

# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO



## FACULTAD DE ECONOMIA

**INFORME DE EXPERIENCIA PROFESIONAL**

**SOBRE:**

**DESESTACIONALIZACION DE SERIES DE TIEMPO ECONOMICAS.  
El caso del Indicador Global de la Actividad Económica (IGAE), México  
(2000-2004).**

**PARA OBTENER EL TITULO DE:**

**LICENCIADO EN ECONOMIA PRESENTA:**

**CARLOS ESPINOSA RAMIREZ**

**DIRECTOR DEL INFORME: MTRO. MIGUEL CERVANTES JIMENEZ**

**MEXICO, D.F. 2006**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**DESESTACIONALIZACION DE SERIES DE TIEMPO ECONOMICAS.  
El caso del Indicador Global de la Actividad Económica (IGAE), México  
(2000-2004).**

<b>INDICE</b>	<b>PAG.</b>
<b>JUSTIFICACIÓN</b> .....	<b>1</b>
<b>OBJETIVO GENERAL Y PARTICULARES</b> .....	<b>5</b>
<i>OBJETIVOS PARTICULARES</i> .....	<b>5</b>
<b>INTRODUCCIÓN</b> .....	<b>6</b>
<b>1. BREVE HISTORIA DEL INSTITUTO NACIONAL DE ESTADÍSTICA GEOGRAFÍA E     INFORMÁTICA (INEGI)</b> .....	<b>9</b>
<b>2. DESESTACIONALIZACIÓN DE SERIES ECONÓMICAS</b> .....	<b>11</b>
2.1. <i>METODOLOGÍA</i> .....	<b>12</b>
<b>3. SELECCIÓN DE LA TRANSFORMACIÓN ESTABILIZADORA DE LA VARIANZA</b> .....	<b>13</b>
<b>4. AJUSTE ESTACIONAL EN X12-ARIMA</b> .....	<b>16</b>
4.1. <i>FASE I AJUSTES PREVIOS DE LA SERIE</i> .....	<b>17</b>
4.2. <i>FASE II AJUSTES INTERMEDIOS</i> .....	<b>21</b>
4.3. <i>FASE III SERIE DESESTACIONALIZADA</i> .....	<b>22</b>
4.4. <i>ESQUEMA SIMPLIFICADO DE DESESTACIONALIZACIÓN CON EL X11</i> .....	<b>23</b>
<b>5. ESTIMACIÓN DE UN MODELO ARIMA</b> .....	<b>24</b>
5.1. <i>CONSIDERACIONES FINALES</i> .....	<b>26</b>
<b>6. CONCEPTO DEL IGAE</b> .....	<b>27</b>
6.1. <i>OBJETIVO DEL IGAE</i> .....	<b>29</b>
6.2. <i>COMPONENTES DEL IGAE</i> .....	<b>30</b>
<b>7. AJUSTES PREVIOS PARA LA DESESTACIONALIZACIÓN</b> .....	<b>33</b>
7.1. <i>IGAE (TOTAL)</i> .....	<b>33</b>
7.2. <i>EFECTOS DE CALENDARIO: SEMANA SANTA, DÍAS LABORABLES Y OUTLIERS</i> .....	<b>36</b>
7.3. <i>IDENTIFICACIÓN DE LA MEJOR DIFERENCIA PARA LA SERIE AJUSTADA DEL IGAE</i> .....	<b>41</b>
<b>8. BÚSQUEDA DEL MEJOR MODELO</b> .....	<b>45</b>
<b>9. DESESTACIONALIZACIÓN DE LA SERIE DEL IGAE</b> .....	<b>55</b>
<b>10. DIAGNÓSTICO DE LA DESESTACIONALIZACIÓN</b> .....	<b>59</b>
10.1. <i>ANÁLISIS DEL MODELO PARA DESESTACIONALIZAR EL IGAE</i> .....	<b>59</b>
<b>11. CONCLUSIONES Y RECOMENDACIONES</b> .....	<b>65</b>
<b>12. ANEXO: SUPUESTOS (BOX-JENKINS)</b> .....	<b>67</b>
<b>BIBLIOGRAFÍA</b> .....	<b>71</b>
<b>ÍNDICE DE TABLAS</b> .....	<b>73</b>
<b>ÍNDICE DE ILUSTRACIONES</b> .....	<b>74</b>

## INTRODUCCIÓN

La necesidad de contar con información económica de coyuntura, veraz y oportuna cada vez es más demandada por los usuarios tanto públicos como privados para realizar distintos estudios, como el desarrollo económico de nuestro país, la situación social de cierta población o toda en conjunto.

La información a que se puede tener acceso son las cifras macroeconómicas, las cuales muestran en forma general la situación de la economía mexicana. De estas cifras podemos hacer uso de las del Producto Interno Bruto, de la Balanza Comercial, del desempleo, de la Actividad Industrial, del Indicador Global de la Actividad Económica principalmente. Esta información se presenta de acuerdo al indicador de que se trate, con una publicación mensual o trimestral.

En México y a nivel internacional, existen diversos demandantes de información oportuna, es por esta razón que existe un organismo encargado de generarla. Esta tarea la desarrolla el Instituto Nacional de Estadística, Geografía e Informática (INEGI), existen otras instituciones oficiales que también calculan y difunden información como el Banco de México (Banxico), la Secretaría de Hacienda y Crédito Público (SHCP) y la Secretaría de Economía (SE), entre otras.

La información recopilada trae consigo efectos ajenos a ella y no propiamente de índole económico como puede ser principalmente la *estacionalidad, entiéndase como una pauta que se repite con periodicidad fija generalmente un año*; ejemplo de ello tenemos las ventas que para el mes de diciembre son considerables, para luego al siguiente caer, este comportamiento de las ventas distorsionan la información y conduce a conclusiones erróneas.

Existen otros efectos que perturban a las series económicas como los llamados efectos de calendario como pueden ser los días laborables y la Semana Santa, estos efectos son de gran importancia ya que pueden influir en que las series no pueden ser comparables entre si por que la composición de los días podría influir de manera desigual en diferentes indicadores o para una misma serie. Por ejemplo, la Semana Santa tiene importancia porque no es una festividad fija, ya que puede ocurrir en marzo o abril. Otras alteraciones que podemos mencionar son los outliers u observaciones aberrantes y valores extremos. Dentro de los

diferentes tipos de outliers podemos encontrar *cambios de nivel, cambios temporales y los aditivos*.

Por los argumentos anteriores y dado que los usuarios de la información requieren cada vez mejor la información, es necesario contar con una metodología que garantice eliminar los efectos no propios de la serie económica de interés.

*Esta tesina de experiencia profesional tiene como objetivo principal desestacionalizar la serie del IGAE por el método X11 y especialmente por el programa X12-ARIMA.* El programa de X12-ARIMA cuenta con diversas técnicas y procesos para lograr el ajuste estacional, así para comprender mejor el funcionamiento de este programa es necesario contar con conocimientos estadísticos, de análisis de series de tiempo y de econometría, entre otros.

La tesina tiene la siguiente estructura. La justificación, los objetivos (particular y generales) y la introducción. En el punto uno presentamos una breve reseña del INEGI, en el punto dos abordamos propiamente el tema que nos compete donde se define a la desestacionalización, así como la metodología que se usará.

En el punto tres se exponen las técnicas para la búsqueda de la transformación de la varianza para seleccionar el modelo a usar, ya sea el multiplicativo o el aditivo.

En el punto cuatro presentamos la parte teórica de los ajustes previos que se realizan en el X12-ARIMA dividido en tres fases; en la primera fase los ajustes previos de la serie. En la fase dos de este punto se presentan los ajustes intermedios que se le hacen a la serie y en la fase tres se recalculan los pasos anteriores para obtener la serie desestacionalizada final, además se presenta un esquema simplificado del funcionamiento del X11, ya que este programa junto con el X12-ARIMA realizan de forma automática la corrección de la serie.

El punto cinco se hace la estimación de un modelo ARIMA que se realiza con la serie ajustada previamente.

En el punto seis se exponen la conceptualización del Indicador Global de la Actividad Económica (IGAE), así como sus componentes y un diagrama del proceso de desestacionalización.

El punto siete trata los ajustes que se realizan a la serie, los efectos de calendario, de semana santa, días laborables y de outliers. Se identifica la mejor diferencia para la serie en estudio, con lo anterior se procede en el punto ocho a estimar el modelo.

En el punto nueve habiendo realizado los ajustes previos a la serie se procede a la desestacionalización de la serie del Indicador Global de la Actividad Económica, así en el punto 10 se realiza un diagnóstico de la serie desestacionalizada y se somete mediante los supuestos de Box y Jenkins a validar el modelo. Para de esta forma discernir si el método utilizado y el paquete X12-ARIMA proporciona una forma apropiada de ajuste estacional para cifras de la economía mexicana; en el punto 11 el trabajo aborda las conclusiones y recomendaciones a que llegamos después de la aplicación de la metodología a una serie económica real, como es el Indicador Global de la Actividad Económica y finalmente el punto 12 contiene los supuestos de Box y Jenkins.

# **1. BREVE HISTORIA DEL INSTITUTO NACIONAL DE ESTADÍSTICA GEOGRAFÍA E INFORMÁTICA (INEGI)**

Con la independencia nacional, la información estadística fue impulsada por el gobierno, con la finalidad de conocer las carencias de la población y las necesidades de infraestructura para impulsar el desarrollo del país. Así para el año de 1882 se funda la Dirección General de Estadística con la finalidad de perfeccionar el ejercicio estadístico para apoyar el diseño, ejecución y evaluación de sus políticas de desarrollo.

Es necesario e importante explicar el porqué el Instituto Nacional de Estadística Geografía e Informática es el encargado de elaborar el Sistema de Cuentas Nacionales para México.

La acción para construir el Sistema de Cuentas Nacionales para México se inició en los años cuarenta, las labores y los productos estaban a cargo por el Banco de México S.A.

En diciembre de 1976, se le responsabilizó en materia de información estadística a la Secretaría de Programación y Presupuesto (SPP); por lo cual en el año de 1977 en esta secretaria se creó la Coordinación General del Sistema Nacional de Información, cuyo objetivo fue construir un nuevo Sistema de Cuentas Nacionales, para lo cual se pidió apoyo a la Organización de las Naciones Unidas (ONU) y a la Comisión Económica para América Latina (CEPAL).

En febrero de 1980, la Coordinación cambió de nombre y paso a ser la Coordinación General de los Servicios Nacionales de Estadística, Geografía e Informática; en abril de 1981 se estableció por decreto presidencial que la (SPP) a través de esta Coordinación era la responsable de elaborar los cálculos macroeconómicos del país.

La necesidad de contar con información económica que facilitara a los distintos agentes económicos conocer con mayor precisión la estructura económica del país permitió la creación del Instituto Nacional de Estadística Geografía e Informática.

Por decreto presidencial, el 25 de enero de 1983 la Coordinación General de los Servicios Nacionales de Estadística, Geografía e Informática pasa a ser el Instituto Nacional de Estadística, Geografía e Informática, dependiente a la entonces Secretaría de Programación y Presupuesto.

Actualmente el INEGI es un órgano desconcentrado de la Secretaría de Hacienda y Crédito Público (SHCP), con autonomía técnica y administrativa, lo que le da la flexibilidad para la consecución de su objetivo básico, que es coordinar y producir los sistemas y servicios nacionales de Estadística y de Información Geográfica de interés nacional.

Su creación fue la respuesta para garantizar la mejora sustancial en la calidad y homogeneidad de la información y además, hacer posible unir esfuerzos de las diferentes instancias y niveles de gobierno, en la integración de un sistema nacional que amplió los alcances que puede tener el uso de la información estadística y geográfica en la instrumentación del plan nacional y de los programas sectoriales y regionales de desarrollo.

Los cambios a nivel internacional como la integración de economías, acuerdos comerciales, y un notable avance tecnológico etc. Propiciaron que los agentes económicos demanden información económica oportuna para una mejor interpretación de las relaciones económicas. La respuesta a esta demanda de información estadística el INEGI creó los Indicadores de Corto Plazo tales como, el Producto Interno Bruto a precios corrientes y constantes, El Indicador Mensual de la Actividad Industrial, la Oferta y Utilización a precios constantes, la Inversión Fija Bruta y el Indicador Global de la Actividad Económica, etc.

La importancia de generar estos indicadores para un mejor análisis económico, es que permiten adentrarse en aspectos estructurales de los fenómenos económicos en un periodo de tiempo muy corto, reflejan su estructura y la interrelación que guardan.

## 2. DESESTACIONALIZACIÓN DE SERIES ECONÓMICAS

Se considera a una serie económica como “*el registro sistematizado de la medición y observación de una numeración efectuada en intervalos de tiempo fijos*”<sup>1</sup>. (En nuestro caso series económicas). Así para describir a la serie económica tiene dos elementos principales las gráficas y las medidas descriptivas para ver la congruencia de los datos.

Existen principalmente dos formas que explican la relación existente entre los componentes de una serie.

La primera es la *multiplicativa*

$$O_t = T_t * S_t * I_t \quad \text{para } t = 1, \dots, N$$

y la segunda *aditiva*

$$O_t = T_t + S_t + I_t \quad \text{para } t = 1, \dots, N$$

En donde  $T_t$ ,  $S_t$  e  $I_t$  son los componentes de tendencia, estacionalidad e irregularidad, respectivamente y  $O_t$  representa a la serie original.

La estacionalidad se describe como “*fluctuaciones periódicas que ocurren cada año, aproximadamente en las mismas fechas y con la misma intensidad las cuales pueden medirse y eliminarse de la serie de tiempo en consideración*”<sup>2</sup>. En particular existen tres características más importantes de los fenómenos estacionales:

- (a) “El fenómeno se repite cada año con cierta regularidad, pero puede evolucionar”.
- (b) “El fenómeno puede medirse y separarse de las otras fuerzas que influyen el movimiento de la serie”.
- (c) “El fenómeno es causado principalmente por fuerzas no-económicas, exógenas al sistema económico y que no pueden controlarse o modificarse por los tomadores de decisiones en el corto plazo”.

*El objetivo de este apartado es exponer los fundamentos de la desestacionalización de series de tiempo económicas y sus componentes, así como un método para ajustar las series.*

---

<sup>1</sup>Guerrero, Víctor (1983a) “Desestacionalización de series de tiempo económicas: parte I. Una introducción a la metodología”, documento de investigación No. 54, Banco de México.

<sup>2</sup>Op. cit. 1983a.

## 2.1. METODOLOGÍA

A la fecha existen diversos métodos para realizar el ajuste estacional de series de tiempo: de regresión, de promedios móviles y basados en modelos. El método que se expondrá en la presente tesina es el de promedios móviles y utiliza el método X11 y su versión más moderna el X12-ARIMA<sup>3</sup>. El problema fundamental en el ajuste estacional de series de tiempo económicas es paradójicamente hacer supuestos “*realistas*” sobre el comportamiento de un fenómeno que no es posible observar directamente.

Los modelos ARIMA buscan representar adecuadamente el comportamiento de los datos observados, la estimación del modelo es sugerida por los datos mismos.

Box y Jenkins consideraron cuatro etapas fundamentales en la revisión de modelos ARIMA.

- 1) Identificación de un posible modelo dentro de la clase de modelos ARIMA, determinación de los valores  $p$ ,  $d$  y  $q$  que especifiquen el modelo ARIMA apropiado para la serie en estudio.
- 2) Estimación de los parámetros involucrados en el modelo, a través de técnicas de estimación no lineal.
- 3) Verificación de que el modelo proporcione un ajuste adecuado y de que los supuestos básicos, implícitos en el modelo, se satisfagan; de no cumplirse los supuestos, se determinan las modificaciones necesarias, mediante la repetición de las etapas anteriores hasta que la verificación indique resultados aceptables.
- 4) Uso del modelo para los fines que el investigador haya tenido en consideración, para un pronóstico, control, simulación o para la explicación de un fenómeno en estudio.

---

<sup>3</sup>Findley, David F. *et al.* (1998). “New capabilities and methods of the X12-ARIMA seasonal.

## JUSTIFICACIÓN

La motivación de presentar el ajuste estacional del Indicador Global de la Actividad Económica como proyecto de experiencia profesional, se atribuye a la importancia del mismo como un indicador de coyuntura mensual, al dar a conocer la tendencia de la economía mexicana. Además de que participé en el cálculo de este indicador en la presentación a precios constantes.

Para la elaboración de este indicador se utilizan los criterios conceptuales, metodológicos y el mismo clasificador de actividades económicas empleado en los cálculos anuales del Sistema de Cuentas Nacionales de México (SCNM). La base del cálculo consiste en elaborar índices de volumen físico de la producción, tipo Laspeyres, para cada uno de los subgrupos de los que se cuenta con información mensual confiable y con la oportunidad requerida para su incorporación. El IGAE tiene su año base en 1993, del cual toma sus ponderaciones y está relacionado con los cálculos mensuales de la Actividad Industrial y el Producto Interno Bruto Trimestral (PIBT).

Tradicionalmente la interpretación de la tasa de crecimiento se hace de un mes respecto al mismo mes del año anterior sin embargo, esto no permite hacer un análisis de corto plazo, con el ajuste estacional el análisis lo podemos realizar mes a mes porque anteriormente los indicadores disponibles no incorporaban ajustes estacionales como en otros países.

Al ser un indicador de coyuntura, es necesario contar con una serie que permita ver mes a mes cuáles son los cambios económicos para la toma de decisiones de los distintos agentes económicos lo cual se obtiene con el ajuste estacional. Dar a conocer este indicador con cifras desestacionalizadas permite con mayor precisión determinar si hay o no cambios importantes que influyan en el comportamiento de la serie. Al comparar las variaciones anuales no se puede precisar el momento en que ocurre una caída o aumento de la actividad económica, en contraste, con la presentación de las series desestacionalizadas se puede observar el comportamiento mes a mes e identificar el movimiento en el margen evitando la distorsión de todo un año. Esta nueva herramienta permite una descripción más exacta del comportamiento de las cifras de los indicadores de coyuntura y en especial del Indicador Global de la Actividad Económica (IGAE).

El INEGI en el año 2000 empezó a publicar las series desestacionalizadas de los indicadores de coyuntura. Esta nueva información permitió un mejor análisis de la actividad económica del país bajo otra perspectiva. El primer indicador de coyuntura que se publicó de manera desestacionalizada fue el Producto Interno Bruto Trimestral. Actualmente la mayoría de los indicadores de coyuntura son desestacionalizados y se publican de acuerdo con un calendario de difusión.

Por lo antes expuesto, considero necesario *describir el proceso que se utiliza para la desestacionalización de series económicas. No sin antes hacer una breve reseña de mi trayectoria laboral dentro del Instituto.*

Tengo 12 años trabajando en el INEGI. Mi primera tarea se desarrolló en la estimación del Producto Interno Bruto Trimestral en el departamento del sector secundario con la responsabilidad de analizar la “división II, cuero, vestido y calzado”, con el objetivo de generar índices de volumen físico para cada producto de cada uno de los subgrupos que conforman la división contenida en la industria manufacturera. Posteriormente pasé al sector primario con la responsabilidad de obtener también índices de volumen físico para la “división I, alimentos, bebidas y tabaco”, con la misma finalidad. Así, mi aportación en esta tarea fue la elaboración de índices de volumen físico tipo Laspeyres cuya fuente primaria es el Indicador Mensual de la Actividad Industrial (IMAI), el cual es un producto de los cálculos del Sistema de Cuentas Nacionales de México (SCNM), de tal forma que *las Cuentas Nacionales son un registro completo y sistematizado de las actividades económicas desarrolladas en un país.*

Además, trabajé en el sector servicios de la misma área (PIBT), colaborando en la Encuesta Trimestral de Servicios para obtener información de subgrupos como los transportes, restaurantes, publicidad y servicios médicos (consultorios, laboratorios y hospitales). Para alcanzar el objetivo anterior fue necesario tener conocimiento de todo el proceso, desde solicitar la información, hacer los oficios para solicitarla de manera formal, colocarlos, hacer llamadas telefónicas para recabar la información y entregar un resultado definitivo de cómo se encuentra cada uno de los indicadores respectivos. Los resultados se presentan en valores constantes para lo cual generé un cuadro con los índices respectivos para su correcta deflatación. También colaboré en la obtención de información de los ferrocarriles, del sector

aéreo, marítimo y de cabotaje, en el cálculo de almacenadoras y automóviles, en obtener un índice de ocupación hotelero desde la petición de la información hasta el procesamiento y entrega de resultados.

Actualmente trabajo en la Coordinación de Análisis Estadístico de Series Económicas, donde se desestacionalizan las series económicas de coyuntura como la Producción Manufacturera y Generación de Electricidad por entidad federativa, el Indicador de la Inversión Fija Bruta, la Balanza Comercial, la Oferta y Demanda de Bienes y Servicios, el Producto Interno Bruto Trimestral a Precios Constantes, los Establecimientos Comerciales y el Indicador Global de la Actividad Económica entre los principales. Este último indicador es el motivo de mi ensayo por presentar una periodicidad mensual y ser oportuno respecto de la tendencia de la economía mexicana. Cabe mencionar que se desestacionalizan regularmente 500 series en el Instituto.

Parte de mis actividades consisten en desestacionalizar el Indicador de la Producción Manufacturera y Generación de Electricidad por entidad federativa, los indicadores principales del Sistema de Indicadores Compuestos Coincidente y Adelantado (SICCA), el total de Minería. Cada uno de los indicadores requiere que cada año al tener la información de un año completo, se le estime un modelo ARIMA. El proceso de obtener un modelo lo detallaré propiamente en el tratamiento de la metodología de ajuste estacional de series económicas.

Cabe señalar que al tener la responsabilidad de desestacionalizar los indicadores de manufacturas que tiene una muestra de 17 entidades federativas, electricidad con 32 entidades, todas y cada una de las series requieren cada año de la estimación de un modelo ARIMA por lo cual les doy seguimiento, desde buscar el modelo apropiado en X12-ARIMA hasta la misma aplicación del modelo para ajustar las series originales.

La periodicidad de los indicadores de coyuntura es mensual a excepción de la Oferta y Demanda Final de Bienes y Servicios y el Producto Interno Bruto que son trimestrales, por tal razón, es necesario tener a la mano un manual que facilite todo el proceso de desestacionalización para el tratamiento de la información para una expedita respuesta y de aceptación.

La experiencia en la desestacionalización de series económicas la adquirí precisamente por la propia necesidad del desarrollo de mis funciones, y observé que existe la posibilidad de

escribir la metodología ya que a la fecha no existe y es necesario, para que no solamente sea un conocimiento del área donde se aplica sino como consulta para todo aquel que se interese en el tema y para el mismo personal de nuevo ingreso. Además de que prácticamente esta nueva forma de analizar una serie económica está poco difundida.

Es conveniente escribir la metodología porque se tendrá a la mano conocimientos básicos de cómo se desestacionaliza propiamente en el INEGI, además de que contribuirá con un valor teórico porque valida los conceptos utilizados.

## **OBJETIVO GENERAL Y PARTICULARES**

*El objetivo general de este ensayo es presentar mi experiencia profesional en el INEGI en la desestacionalización de series económicas de coyuntura, mismas que son útiles para la toma de decisiones en el corto plazo de los distintos agentes económicos, particularmente en la desestacionalización del IGAE.*

### **OBJETIVOS PARTICULARES**

- 1. Describir el proceso que se utiliza para desestacionalizar el Indicador Global de la Actividad Económica.*
- 2. Determinar si la desestacionalización de series económicas como el IGAE permite una mayor precisión en el análisis del comportamiento oportuno de la tendencia de la economía mexicana.*
- 3. Evaluar cuales de sus componentes (Sector Agropecuario, Sector Industrial y Sector Servicios) inciden más en el comportamiento del IGAE.*

### 3. SELECCIÓN DE LA TRANSFORMACIÓN ESTABILIZADORA DE LA VARIANZA

El objetivo de este apartado es seleccionar el modelo multiplicativo o aditivo mediante la transformación potencia más adecuada para estabilizar la varianza de la serie. Esto con la finalidad de volver estacionaria la serie, mediante la aplicación del operador diferencia para evitar una posible tendencia polinomial de orden  $d$ .

Para discriminar objetivamente entre modelos multiplicativos y aditivos se sigue el método propuesto por Guerrero<sup>4</sup>, el cual consiste en encontrar la transformación potencia más adecuada para estabilizar la varianza de la serie. El método sugiere elegir la potencia  $\lambda$  de tal manera que satisfaga la siguiente relación.

$$(1) \quad \frac{\sigma_t}{\mu_t^{1-\lambda}} = K \quad \text{para } t = 1, K, n$$

En donde  $\sigma_t$  representa la desviación estándar y  $\mu_t$  la media de la serie original,  $n$  el número de observaciones que tiene la serie  $O_t$  y  $K$  es una constante de proporcionalidad.

La relación entre la transformación potencia más adecuada a la serie y la forma funcional de los componentes de la serie original es muy estrecha, debido a que *es posible expresar un modelo multiplicativo de manera equivalente a un aditivo*.

$$\log O_t = \log T_t + \log S_t + \log I_t \quad \text{para } t = 1, \dots, N$$

Entonces se utilizaría el modelo aditivo y el problema se reduce a discriminar entre emplear el logaritmo de las observaciones o las observaciones originales. Si la amplitud del componente estacional es proporcional (constante) respecto a la tendencia media anual, el modelo será multiplicativo (aditivo). Así, si se supone que la variación en grupos anuales de datos es causada esencialmente por las fluctuaciones estacionales, se puede obtener una relación del tipo.

---

<sup>4</sup> Guerrero, Víctor; "Análisis estadístico de series de tiempo económicas", México, 2ª. Edición. México, D.F. 2003, ed. Thomson.

$$(2) \quad \frac{\sigma_t}{\mu_t} = K \text{ para } t = 1, K, N$$

o bien

$$(3) \quad \sigma_t = K \text{ para } t = 1, K, N$$

Esto exige emplear un modelo multiplicativo en el caso (2) y uno aditivo en el caso (3).

Desde luego ésta es una aproximación, pero de esta manera se evita el tener que estimar una tendencia previa de la serie, ya que el supuesto en el que se basa la relación anterior es que la tendencia crece de un año a otro, pero permanece constante en cada año.

Por otro lado, el procedimiento propuesto por Guerrero se justifica con el argumento de que una relación más general como (1), contiene las expresiones (2) y (3). Así en la ecuación (1) si la potencia es  $\lambda = 0$ , se obtiene (2) y si la potencia es  $\lambda = 1$ , entonces se tiene la ecuación (3). Lo importante en la ecuación (1) es que demuestra qué surge de la búsqueda de una transformación estabilizadora de la varianza. Además si (1) se satisface razonablemente para alguna potencia  $[\lambda]$ , sería adecuado utilizar la transformación potencia.

$$O'_t = O_t^\lambda \quad \text{si } \lambda \neq 1,$$

$$O'_t = \log O_t \quad \text{si } \lambda = 0.$$

Es importante señalar que los modelos imponen restricciones a su uso, en especial, el multiplicativo no puede emplearse para una serie que contenga ceros o datos negativos, de igual manera que no puede aplicarse el logaritmo de cifras negativas; así pues, para series con datos negativos o ceros es necesario usar el modelo aditivo. *Sin embargo, la transformación potencia más adecuada para la serie decidirá que modelo de descomposición utilizar.*

El método funciona de la siguiente manera: divídase a las  $N$  observaciones de la serie en  $H$  grupos que contengan  $R = (N - n)/H$  observaciones contiguas cada uno, dejando fuera de los cálculos a un total de  $n$  observaciones ( $0 \leq I \leq R$ ) ya sean del principio o al final de la serie; lo importante es cuidar que exista homogeneidad entre los grupos y que todos ellos contengan el mismo número de observaciones, para estimar la media y la desviación estándar dentro de

cada grupo; de esta forma se tendrán  $H$  parejas de valores comparables  $\{\sigma_h, \bar{O}_h\}$ ,  $h = 1, \dots, H$ , con los cuales podrá construirse la siguiente tabla:

Tabla 3.1. México: Cálculos para seleccionar la potencia ( $\lambda$ ) de la Transformación estabilizadora de la varianza.

Potencia ( $\lambda$ )					
Grupos	-1	-0.5	0	0.5	1
1	$\sigma_1/\bar{O}_1^2$	$\sigma_1/\bar{O}_1^{1.5}$	$\sigma_h/\bar{O}_1$	$\sigma_h/\bar{O}_1^{0.5}$	$\sigma_1$
2	$\sigma_2/\bar{O}_2^2$	$\sigma_2/\bar{O}_2^{1.5}$	$\sigma_h/\bar{O}_2$	$\sigma_h/\bar{O}_2^{0.5}$	$\sigma_2$
h	$\sigma_h/\bar{O}_h^2$	$\sigma_h/\bar{O}_h^{1.5}$	$\sigma_h/\bar{O}_h$	$\sigma_h/\bar{O}_h^{0.5}$	$\sigma_h$
...					
H	$\sigma_H/\bar{O}_H^2$	$\sigma_H/\bar{O}_H^{1.5}$	$\sigma_h/\bar{O}_H$	$\sigma_h/\bar{O}_H^{0.5}$	$\sigma_H$
Coefficiente de variación	CV(-1)	CV(-0.5)	CV(0)	CV(0.5)	CV(1)

En donde, si  $O_{h,r}$  es la  $r$ -ésima observación del grupo  $H$ ,

$$(4) \quad \bar{O}_h = \frac{\sum_{r=1}^R O_{h,r}}{R}, \quad \sigma_h = \sqrt{\frac{\sum_{r=1}^R (O_{h,r} - \bar{O}_h)^2}{(R-1)}}.$$

Además, el coeficiente de variación se calcula como la desviación estándar de ( $\lambda$ ), entre la media  $M(\lambda)$ , es decir,

$$(5) \quad CV(\lambda) = \frac{de(\lambda)}{M(\lambda)} \quad \text{con} \quad M(\lambda) = \sum_{h=1}^H \frac{\sigma_h/\bar{O}_h^{1-\lambda}}{H} \quad \text{y}$$

$$de(\lambda) = \sqrt{\frac{\sum_{h=1}^H \left[ \left( \frac{\sigma_h}{\bar{O}_h^{1-\lambda}} \right) - M(\lambda) \right]^2}{H-1}},$$

Entonces, ya que se desea satisfacer la ecuación (1) con los valores muestrales, deberá elegirse la potencia que proporcione el mínimo coeficiente de variación, debido a que ese valor satisfaga de manera aproximada la relación.

$$\frac{S_h}{\bar{O}_h^{1-\lambda}} = K, \quad \text{para } h = 1, \dots, H.$$

## 4. AJUSTE ESTACIONAL EN X12-ARIMA

En este paso se estiman componentes, como son los valores extremos y los efectos de calendario, es decir, que algunas festividades estén fijas para determinadas fechas, que los meses tengan más o menos días o que se presenten más o menos fines de semana.

*El objetivo de este apartado es presentar los ajustes previos que se realizan por la distinta composición de los días de la semana, de Semana Santa y de valores extremos así como el algoritmo que a grosso modo ejecuta el programa X12-ARIMA.*

Los efectos de calendario<sup>5</sup> más comunes son: la distinta composición de los días de la semana y la Semana Santa. La importancia de eliminar estos efectos de la serie original es que se consideran factores determinísticos de la serie original y por tanto, no provienen del proceso estocástico que generó la serie original. También es necesario eliminar aquellos valores extremos que se deben a causas ajenas al proceso estocástico generador de la serie original. Ambos ajustes son necesarios para estimar el modelo ARIMA de la serie en estudio, para poder obtener pronósticos con error cuadrático medio mínimo.

Los efectos por la distinta composición de los días de la semana se estiman por métodos de regresión en el componente irregular de la serie, por lo que es necesaria una estimación previa de este componente. Debido a que el componente irregular o residual se supone que no está auto-correlacionado, es un buen candidato para aplicar mínimos cuadrados ordinarios (MCO) para estimar componentes adicionales como son los efectos de calendario.

Los valores extremos se definen como aquellos mayores, en valor absoluto, a 2.5 desviaciones estándar ( $\sigma$ ).

*En la descripción del ajuste estacional la descomposición de la serie puede ser representada por un proceso aditivo, multiplicativo o el logaritmo de cualquiera de los dos procesos; se realizan ajustes previos por la distinta composición de los días de la semana (Trading Day), de Semana Santa (Easter) y de valores extremos (Outliers) con límite superior de 2.5  $\sigma$  para éste último.*

---

<sup>5</sup> Para mayor detalle de estos efectos ver a: Guerrero, Víctor (1983b) Desestacionalización de series de tiempo económicas”, parte II Ajustes previos a la desestacionalización, documento de investigación No. 55 Banco de México.

El algoritmo que el programa X12-ARIMA ejecuta se divide en tres fases como a continuación se presenta:

#### 4.1. FASE I AJUSTES PREVIOS DE LA SERIE

(a) Estimación inicial de la tendencia por medio de un promedio móvil centrado de 12 términos en la serie original.

$$T_t^{(1)} = \frac{1}{24}O_{t-6} + \frac{1}{12}O_{t-5} + \dots + \frac{1}{12}O_t + \dots + \frac{1}{12}O_{t+5} + \frac{1}{24}O_{t+6}.$$

(b) Calcular la razón SI (Seasonal Irregular) inicial.

$$SI_t^{(1)} = \frac{O_t}{T_t^{(1)}}.$$

(c) Obtener el factor estacional inicial con una media móvil estacional de  $3 \times 3$ .

$$\hat{S}_t^{(1)} = \frac{1}{9}SI_{t-24}^{(1)} + \frac{2}{9}SI_{t-12}^{(1)} + \frac{3}{9}SI_t^{(1)} + \frac{2}{9}SI_{t+12}^{(1)} + \frac{1}{9}SI_{t+24}^{(1)}.$$

(d) Factor estacional inicial.

$$S_t^{(1)} = \frac{\hat{S}_t^{(1)}}{\frac{1}{24}\hat{S}_{t-6}^{(1)} + \frac{1}{12}\hat{S}_{t-5}^{(1)} + \dots + \frac{1}{12}\hat{S}_t^{(1)} + \dots + \frac{1}{12}\hat{S}_{t+5}^{(1)} + \frac{1}{24}\hat{S}_{t+6}^{(1)}}.$$

(e) Ajuste por valores extremos.

(e.1) Obtener la serie irregular inicial.

$$I_t^{(1)} = \frac{SI_t^{(1)}}{S_t^{(1)}}, \text{ donde } \bar{I}_t^{(1)} = \frac{1}{T} \sum_{t=1}^T I_t^{(1)} = 1.$$

(e.2) Calcular las desviaciones estándar móviles de 5 años de  $I_t^{(1)}$  y eliminar las observaciones en el componente irregular inicial de los años centrales ( $3 \leq n \leq N - 2$ ) que sean mayores a  $2.5 \sigma$ . Recalcular las desviaciones estándar móviles de 5 años. Observar cuales meses en  $I_t^{(1)}$  son menores a  $1.5 \sigma$ , cuales mayores a  $2.5 \sigma$  y cuales se encuentran entre  $1.5 \sigma$  y  $2.5 \sigma$ .  $N$  indica el número total de años que contiene la serie y debe ser mayor o igual a 5,  $n$  es un índice que indica el año y  $t$  el mes del año  $n$  ( $1 \leq t \leq 12$ ).

$$\sigma_1, \sigma_2 = \sigma_3$$

$$\sigma_3 = \sqrt{\frac{1}{60} \sum_{t=1}^{60} (I_t^{(1)} - 1)^2}$$

$$\sigma_4 = \sqrt{\frac{1}{60} \sum_{t=13}^{72} (I_t^{(1)} - 1)^2}$$

$$\sigma_n = \sqrt{\frac{1}{60} \sum_{t=12(n-3)+1}^{12(n+2)} (I_t^{(1)} - 1)^2}$$

$$\sigma_{N-2} = \sqrt{\frac{1}{60} \sum_{t=12(N-5)+1}^{12N} (I_t^{(1)} - 1)^2}$$

$$\sigma_{N-1}, \sigma_N = \sigma_{N-2}$$

**(e.3)** Se asignan ponderaciones  $w_t$  (para cada mes  $t$  del año  $n$ ) las cuales están entre 0 y 1 a la serie  $I_t^{(1)}$  usando sus respectivos errores estándar  $\sigma_n$  de acuerdo con la siguiente regla<sup>6</sup>.

Defínase en primer instancia  $\sigma_t$  para  $1 \leq t \leq 12N$  como el error estándar para cada mes  $t$ , determinado por su año  $n$ , el cual satisface que  $12(n-1) + 1 \leq t \leq 12n$  y sea  $\sigma_t = \sigma_n$

$$w_t = 1 \text{ si } |I_t^{(1)} - 1| \leq 1.5\sigma_t.$$

$$w_t = \frac{(2.5\sigma_t - |I_t^{(1)} - 1|)}{2.5\sigma_t - 1.5\sigma_t} \text{ si } 1.5\sigma_t < |I_t^{(1)} - 1| < 2.5\sigma_t.$$

$$w_t = 0 \text{ si } |I_t^{(1)} - 1| \geq 2.5\sigma_t.$$

**(e.4)** Reemplazo de los valores  $SI_t$ . Dadas las series  $SI_t^{(1)}$  y  $S_t^{(1)}$  calculadas en (b) y (d)

Cuando  $w < 1$ , si  $w_{t \pm 12} = w_{t \pm 24} = 1$ , reemplazar  $SI_t^{(1)}$  por

$$SI_{t,w}^{(1)} = \frac{1}{4 + w_t} [SI_{t-24}^{(1)} + SI_{t-12}^{(1)} + w_t SI_t^{(1)} + SI_{t+12}^{(1)} + SI_{t+24}^{(1)}].$$

Si  $w_t = 1$ , no hacer reemplazo alguno en ese mes.

**(f)** Para ajustar la serie original por días laborables, obtener el componente irregular ajustado por valores extremos.

---

<sup>6</sup>Recuérdese que para las primeras 36 observaciones se usa  $\sigma_3$ , para las observaciones 37 a 48  $\sigma_4$ , de las observaciones 49 a 60  $\sigma_5$  y así sucesivamente.

$$I_{t,w}^{(1)} = \frac{SI_{t,w}^{(1)}}{S_{t,w}^{(1)}}$$

(f.1) Estimar por MCO la siguiente regresión.

$$\left(I_{t,w}^{(1)} - 1\right) = \frac{\sum_{j=1}^7 X_{jt} \beta_j}{N_t} + e_t,$$

Donde  $X_{jt}$  es el número de veces que se presenta el día  $j^7$  en el mes  $t$ ,  $N_t$  es el número de días que tiene el mes  $t^8$  y  $e$  es un término de error.

(f.2) Sea  $b_j$  la estimación de  $\beta_j$  hacer una prueba de significancia global de la regresión. Si las  $b_j$  son estadísticamente distintas de cero, realizar el ajuste. Si no lo son, entonces omitir este paso y continuar en (g).

(f.3) Calcular los factores de días laborables  $M_t$  de la siguiente manera:

$$M_t = \frac{\sum_{j=1}^7 X_{jt} (b_j - 1)}{N_t}.$$

(f.4) Ajustar la serie original por los efectos de Semana Santa.

En este modelo se asume que el efecto de la Semana Santa se incrementa o decrece linealmente durante los  $k$  días,

$f(Z_i) = 1$	si $Z_i \leq 9$ (la Semana Santa cae en marzo)
$f(Z_i) = \frac{k + 9 - Z_i}{k}$	si $9 \leq Z_i \leq k + 9$ (la Semana Santa cae en abril y antes del día $k$ de abril)
$f(Z_i) = 0$	si $Z_i \geq 9 + k$ (la Semana Santa cae en o después del día $k$ de abril)

Donde  $Z_i$  es el número de días entre el domingo de pascua y el 22 de marzo del año  $i$ .

La estimación de los efectos de Semana Santa es como sigue:

$$\hat{E}_i = \frac{1}{2} f(Z_t) \left[ \frac{\sum_{i \in M} (I_{i,j+1} - I_{i,j})}{n_M} - \frac{\sum_{i \in A} (I_{i,j+1} - I_{i,j})}{n_A} \right]$$

<sup>7</sup> Lunes  $j=1$ , martes  $j=2, \dots$ , domingo  $j=7$ .

<sup>8</sup>  $N_t$  solo puede tener los siguientes valores 31, 30, o 28, 29 en febrero.

Donde  $I_{i,j}$  es el componente irregular ajustado por valores extremos  $[I_{t,w}^{(2)}]$ ,  $E_i$  es el efecto de Semana Santa,  $i$  denota el año,  $j$  marzo,  $j + 1$  abril,  $n_M$  es el número de años donde la Semana Santa cae en marzo,  $n_A$  el número de años donde cae en o después del  $k$ -ésimo día de abril,  $M$  es el subconjunto de años en los que cae en marzo y  $A$  el subconjunto de años donde sucede en o después del  $k$ -ésimo día de abril.

Los posibles valores de  $k$  son de 1 a 9 días. X12-ARIMA puede estimar el valor de  $k$  el cual selecciona el valor de  $k$  que minimiza la suma de errores al cuadrado entre la serie observada y la estimada. Los factores de Semana Santa se calculan como sigue:

$$SS_t = \begin{cases} \text{Factor en marzo}_i = 1 - \hat{E}_i \\ \text{Factor en abril}_i = 1 + \hat{E}_i \end{cases}$$

Donde,  $SS_t$  es el factor de Semana Santa.

Para probar si el efecto mencionado está presente en la serie se hace una prueba de significancia. Para llevarla a cabo, sea:

$$\nabla I_i = I_{i,j+1} - I_{i,j},$$

$$\overline{\nabla I}_M = \frac{1}{n_M} \sum_{i \in M} \nabla I_i,$$

$$\overline{\nabla I}_A = \frac{1}{n_A} \sum_{i \in A} \nabla I_i \text{ y}$$

$$\overline{\nabla I} = \frac{1}{n_M + n_A} \sum_{\substack{i \in A \\ i \in M}} \nabla I_i.$$

Posteriormente se calcula un estadístico  $\mathfrak{F}$ , el cual se compara con el valor de una  $F$  con un grado de libertad en el numerador y  $(n_M + n_A - 2)$  en el denominador. Rechazar la hipótesis nula de que los efectos de Semana Santa no están presentes en la serie, si el valor en tablas de la  $F$  es menor que el estadístico  $\mathfrak{F}$ .

$$\mathfrak{F} = \frac{1}{n_M + n_A - 2} \left\{ \frac{n_M (\overline{\nabla I}_M - \overline{\nabla I})^2 + n_A (\overline{\nabla I}_A - \overline{\nabla I})^2}{\sum_{\substack{i \in M \\ i \in A}} (\nabla I_i - \overline{\nabla I})^2 - [n_M (\overline{\nabla I}_M - \overline{\nabla I})^2 + n_A (\overline{\nabla I}_A - \overline{\nabla I})^2]} \right\}.$$

(f.5) Se divide la serie original entre el producto de los factores de días laborables y Semana Santa estimados en los pasos anteriores.

$$O_t^{dl} = \frac{O_t}{M_t S S_t}.$$

Usando  $SI_{t,w}^{(1)}$  recalcular (c) y (d) para obtener la serie desestacionalizada inicial.

$$A_t^{(1)} = \frac{O_t^{dl}}{S_{t,w}^{(1)}}.$$

## 4.2. FASE II AJUSTES INTERMEDIOS

Obtener una tendencia intermedia por medio del filtro lineal de Henderson<sup>9</sup> para  $H = 6$ .

$$T_t^{(2)} = \sum_{j=-H}^H h_j^{(2H+1)} A_{t+j}^{(1)}.$$

Posteriormente se calculan los factores  $SI_t$  intermedios.

$$SI_t^{(2)} = \frac{O_t^{dl}}{T_t^{(2)}}$$

El factor estacional preliminar se obtiene usando un promedio móvil estacional de  $3 \times 5$  como en la fase anterior.

$$S_t^{(2)} = \frac{1}{15} SI_{t-36}^{(2)} + \frac{2}{15} SI_{t-24}^{(2)} + \frac{3}{15} SI_{t-12}^{(2)} + \frac{3}{15} SI_t^{(2)} + \frac{3}{15} SI_{t+12}^{(2)} + \frac{2}{15} SI_{t+24}^{(2)} + \frac{1}{15} SI_{t+36}^{(2)}.$$

El factor estacional intermedio se calcula de la siguiente forma.

$$S_t^{(2)} = \frac{\hat{S}_t^{(2)}}{\frac{1}{24} S_{t-6}^{(2)} + \frac{1}{12} S_{t-5}^{(2)} + \Lambda + \frac{1}{12} S_t^{(2)} + \Lambda + \frac{1}{12} S_{t+5}^{(2)} + \frac{1}{24} S_{t+6}^{(2)}}.$$

Para obtener la serie desestacionalizada intermedia.

---

<sup>9</sup> Este filtro se conoce en inglés como *low-pass linear filter*, el cual remueve de la serie el componente de corto plazo de ésta. También se puede decir que este filtro suprime el componente que está en las frecuencias más altas de la función de densidad espectral de una serie estacionaria y esencialmente preserva las magnitudes de los componentes cuya frecuencia se encuentre cercana a cero. De manera similar se espera este resultado para series no estacionarias [ver Findley *et al.* (1998)].

$$W A_t^{(2)} = \frac{O_t^{dl}}{S_t^{(2)}}.$$

### 4.3. FASE III SERIE DESESTACIONALIZADA

Se vuelven a calcular los pasos anteriores usando como original la serie ajustada por valores extremos y días laborables, la cual se obtiene de la siguiente forma:

$$O_t^{(2)} = A_t^{(2)} S_t^{(2)}.$$

En el paso (a) de la fase II, el programa X12-ARIMA decidirá cuales de los tres posibles filtros de Henderson (H = 4, 6 u 11) aplicará. Para elegir el filtro, se calcula el siguiente cociente:

$$I_t^{H6} = \frac{A_t^{(2)}}{T_t^{(2)}}.$$

Después se calcula la siguiente razón:

$$\frac{\bar{I}}{\bar{C}} = \frac{\frac{\sum_{t=8}^{T-6} |I_t^{H6} - I_{t-1}^{H6}|}{I_{t-1}^{H6}}}{\frac{(T-13) \sum_{t=8}^{T-6} |T_t^{H6} - T_{t-1}^{H6}|}{T_{t-1}^{H6}}}.$$

La regla de decisión para elegir el  $H$  se puede ver en la tabla siguiente.

Tabla 4.1. Decisión entre diferentes filtros de Henderson.

Low noise	$\bar{I}/\bar{C} \leq 1$	H = 4
Medium noise	$1 \leq \bar{I}/\bar{C} \leq 3$	H = 6
High noise	$\bar{I}/\bar{C} \geq 3.5$	H = 11

Seleccionado el valor de  $H$  se calculan los pasos desde (a) hasta (e) de la fase II. Donde  $A_t^{(2)}$  será la serie desestacionalizada final.

Para obtener los factores estacionales pronosticados un año hacia adelante, X12-ARIMA los calcula como sigue:

$$S_{j,t+1} = S_{j,t} + \frac{1}{2}(S_{j,t} - S_{j,t-1}) \quad \text{donde } j \text{ es el mes del año } t .$$

#### 4.4. ESQUEMA SIMPLIFICADO DE DESESTACIONALIZACIÓN CON EL X11<sup>10</sup>

<p><b>ETAPA A: Ajustes previos</b></p> <ul style="list-style-type: none"> <li>-Para efectos conocidos e importantes</li> <li>-Para días hábiles</li> </ul>
<p><b>ETAPA B: Primera corrección automática de la serie</b></p> <ul style="list-style-type: none"> <li>-Estimación de la componente irregular</li> <li>-Detección y corrección automática de los puntos atípicos</li> <li>-Corrección de los efectos de días hábiles</li> </ul>
<p><b>ETAPA C: Segunda corrección automática de la serie</b></p> <ul style="list-style-type: none"> <li>-Estimación de la componente irregular</li> <li>-Detección y corrección automática de los puntos atípicos</li> <li>-Corrección de los efectos de días hábiles</li> </ul>
<p><b>ETAPA D: Desestacionalización</b></p> <ul style="list-style-type: none"> <li>-Cálculo de la serie desestacionalizada provisoria (tabla D1 a D6)</li> <li>-Alisado de la serie desestacionalizada con una media móvil de Henderson y nueva estimación de los coeficientes estacionales (tabla D7 a D10)</li> <li>-Cálculo de la serie desestacionalizada definitiva (tabla D11), extracción de la componente tendencia-ciclo (tabla D12) y de la componente irregular (tabla D13)</li> </ul>
<p><b>ETAPA E: Componentes corregidos de los valores muy atípicos</b></p>
<p><b>ETAPA F: Evaluación de la calidad de la desestacionalización</b></p>
<p><b>ETAPA G: Gráficos</b></p>

X12-ARIMA se basa en este principio pero además propone un módulo muy completo (Reg-ARIMA) que permite corregir la serie inicial de todo tipo de efectos indeseables. Reg-ARIMA esta en X12-ARIMA que de forma *automática* además de eliminar efectos indeseables se obtiene el *easter*, *trading day* y los *outliers* (aditivos, cambio de nivel y el cambio temporal).

<sup>10</sup> Dominique Ladiray, Benoît Quenneville “Desestacionalizar con el Método X11”; Prefacio de Allan Young; Traducción al castellano: Eduardo Crivisqui 2000 – 2001, Documento en PDF.

## 5. ESTIMACIÓN DE UN MODELO ARIMA

La estimación consta de tres partes: la **primera** es la identificación de un modelo ARIMA utilizando las funciones de auto correlación y de auto correlación parcial. En la **segunda** parte, una vez identificado un modelo ARIMA  $(p,d,q) \times (P,D,Q)E$ , se estiman los parámetros por medio del método de máxima verosimilitud. La **tercera** parte se le conoce como verificación, en donde se analiza la serie de residuales del modelo estimado para saber si existe evidencia estadística de alguna dependencia de esta serie, es decir que sea ruido blanco y se verifican ciertos supuestos.

*Este apartado se encarga de manifestar que la estimación del modelo ARIMA tiene como propósito pronosticar la serie para evitar el inconveniente de perder datos tanto al inicio como al final de la serie debido a la aplicación de medias móviles.*

El programa X12-ARIMA tiene una opción automática, la cual tiene cuatro modelos ARIMA básicos<sup>11</sup> y toma el primero de ellos que haya satisfecho los supuestos que el programa verifica<sup>12</sup>. Es importante señalar que es necesario usar un modelo ARIMA para desestacionalizar la serie, pero en ocasiones los modelos automáticos no satisfacen todos los supuestos. De esta forma, la opción automática se utiliza sólo como una primera aproximación al mejor modelo ARIMA de la serie. A continuación se detalla la parte de estimación y verificación, de un modelo ARIMA, los supuestos pueden ser consultados en el anexo de este trabajo.

Sea  $z_t$  la serie original ajustada previamente por valores extremos, días festivos, Semana Santa y días laborables  $[O_t^{(2)}]$ , la cual tiene  $N = n + d$  observaciones y sigue un proceso ARIMA(p,d,q)<sup>13</sup>. De estas observaciones es posible generar una nueva serie  $w_t$  con  $n = N - d$  observaciones y donde  $w_t = \nabla^d z_t$ <sup>14</sup>. Por lo que ajustar un modelo ARIMA para  $\nabla^d z_t$  es Equivalente a estimar un modelo ARMA  $(p,q)$  para  $w_t$ . El modelo a estimar puede escribirse como:

---

<sup>11</sup> Utilizando la notación de Box y Jenkins, los cuatro modelos son: a) (0,1,1) (0,1,1)E; b) (0,1,2) (0,1,1)E; c) (2,1,0) (0,1,1)E y d) (0,2,2) (0,1,1)E. Donde E es la periodicidad estacional de la serie, 12 para mensuales y 4 para trimestrales.

<sup>12</sup> Estos son: que el error cuadrático medio del pronóstico sea menor a 15% y que el nivel de significancia de la distribución  $\chi^2$  de la prueba de aleatoriedad en los residuales de Ljung y Box sea mayor o igual a 5%.

<sup>13</sup> Por el momento se omitirá la parte estacional por simplicidad en la notación.

<sup>14</sup>  $\nabla$  indica un operador diferencia,  $\nabla^d = (1 - z_{t-1})^d$ .

(6)	$a_t = w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$
-----	---

Donde se supone que el valor esperado de  $w_t$  es igual a cero, los primeros  $p$  valores de  $w$  y los primeros  $q$  valores de  $a$  están dados, los cuales se denotan como  $w^*$  y  $a^*$ . Por lo tanto, es posible calcular la secuencia de valores de  $a_t$  ( $\phi, \theta \mid w^*, a^*$ ) para  $t = 1, 2, \dots, n$ . Adicionalmente se supondrá que  $(a_t \sim N(0, \sigma_a^2))$  son independientes, idénticamente distribuidas. Por lo tanto, la función de densidad de los errores es:

(7)	$f(a_1, \dots, a_n) \propto \sigma^{-n} \exp \left[ - \left( \sum_{t=1}^n \frac{a_t^2}{2\sigma_a^2} \right) \right],$
-----	---

Por lo que la función de verosimilitud de la serie  $w_t$  se puede expresar como:

(8)	$L(\phi, \theta, \sigma_a) = -n \ln \sigma_a - \frac{\sum_{t=1}^n a_t^2(\phi, \theta \mid w^*, a^*, w)}{2\sigma_a^2}.$
-----	--

Maximizando la ecuación (8) se obtienen los estimadores máximo verosímiles de  $\phi, \theta$  y  $\sigma_a$ . Estos estimadores son BAN (**best asymptotically normal**), maximizar (8) es equivalente a minimizar la suma de los residuales al cuadrado, por lo que es equivalente a estimar por mínimos cuadrados condicionados en  $w^*$  y en  $a^*$ .

Este método exige estimar valores iniciales para  $w$  y para  $a$  ( $w^*, a^*$ ). Una forma de estimarlos es cuando el tamaño de la muestra es grande o moderada y hacerlos igual a su valor esperado incondicional.

La esperanza incondicional de  $a^*$  es cero y si el modelo no contiene alguna parte determinística y  $E[w] = 0$ , el valor esperado incondicional de  $w^*$  será también cero. En caso de que alguna de las raíces del polinomio auto regresivo sea cercano a 1, entonces la esperanza incondicional de  $w^*$  no será necesariamente cero, pues el proceso es aproximadamente no estacionario (en sentido débil). Sin embargo, para series estacionales no es recomendable estimar los parámetros por este método, por lo que la estimación por máxima verosimilitud exacta es aconsejable.

La diferencia entre estimar por máxima verosimilitud condicional o exacta es que la segunda no restringe la función de verosimilitud en  $w^*$  y  $a^*$ .

*La verificación de los supuestos del modelo ARIMA*<sup>15</sup> estimado se concentra principalmente en los residuales estimados. La idea de verificar un modelo se sustenta en que todo modelo es erróneo, ya que los modelos son sólo representaciones simplificadas de la realidad. Por lo que si se proponen varios modelos, se elegirá aquel que presente menos fallas, o las menos importantes. Por lo tanto habrá que poner a prueba todos los posibles modelos para detectar sus fallas, que se miden como violaciones a los supuestos que fundamentan al modelo.

## **5.1. CONSIDERACIONES FINALES**

Para evitar revisiones frecuentes en la serie desestacionalizada se estima un modelo ARIMA cada año. Además, una vez estimado un modelo ARIMA, se fijan para todo el año los parámetros estimados. Algunas revisiones se hacen durante el año, pero no son muy frecuentes. El tipo de transformación estabilizadora de la varianza de la serie y el tipo de descomposición también se analiza cada vez que se completa un año.

Es importante señalar que el programa X12-ARIMA estima cuatro modelos ARIMA como se mencionó con anterioridad, a este módulo del programa se le llama “opción automática” sin embargo, solo se usa para darse una primera idea del modelo ARIMA que sigue la serie.

La metodología mencionada se aplicará a una serie económica real con la finalidad de desestacionalizarla. Por tal razón consideremos al Indicador Global de la Actividad Económica como la serie apropiada para tal objetivo además de ser un indicador oportuno de evolución de la economía mexicana. En el punto 6 presentamos la conceptualización del IGAE.

---

<sup>15</sup> Recuérdese que la verificación de los supuestos es indispensable para considerar un modelo como apropiado, ver a Guerrero, Víctor; “Análisis estadístico de series de tiempo económicas”, 2ª. Edición. Cap.IV. México, D.F. 2003, ed. Tomson.

## 6. CONCEPTO DEL IGAE

El Indicador Global de la Actividad Económica (IGAE), muestra la evolución de la actividad económica del país, con periodicidad mensual y una oportunidad prevista *entre 55 y 57 días después de concluido el mes de referencia*<sup>16</sup>.

*El objetivo de este capítulo es mostrar como se conceptualiza el IGAE, sus fuentes de información y componentes, presentando de forma práctica como se estructura el Indicador Global de la Actividad Económica.*

Para la cuantificación del Indicador Global de la Actividad Económica se utiliza el esquema conceptual y metodológico del Sistema de Cuentas Nacionales de México (SCNM).

Así, el IGAE se expresa mediante un índice de volumen físico base 1993=100 tipo Laspeyres<sup>17</sup>. También emplea la misma clasificación por actividades económicas y fuentes básicas de información con oportunidad mensual.

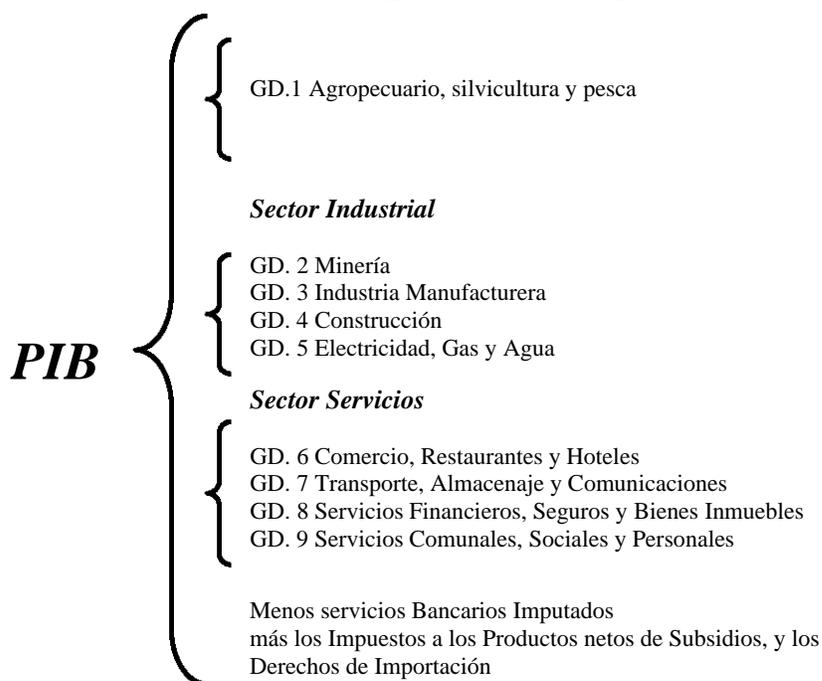
Es importante mencionar que la información básica que incorpora el IGAE es preliminar y está sujeta a revisión por parte de las empresas y organismos públicos y privados; asimismo, no incluye a todas las ramas de actividades económicas como lo hace el Producto Interno Bruto Trimestral (PIBT), el cual considera a las nueve grandes divisiones y las divisiones de la Industria Manufacturera, menos los cargos por los servicios bancarios imputados, más los impuestos a los productos netos de subsidios y los derechos de importación, todo ello suma la totalidad del PIB igual al 100%. La ilustración siguiente simplifica la composición del Producto Interno Bruto.

---

<sup>16</sup> INEGI; (2000) “Sistema de Cuentas Nacionales de México”, Indicador Global de la Actividad Económica (Mensual).

<sup>17</sup> Para mayor detalle sobre índices ver a: Núñez del Prado Benavente Arturo; estadística Básica para Planificación; 16ª. Ed. Siglo XXI.

### Ilustración 6.1. México: Composición del PIB por Grandes Divisiones.



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Por lo anterior, los resultados del IGAE pueden diferir de los del PIBT y debe considerársele como un indicador de la evolución o dirección de la actividad económica en el país a corto plazo.

Las fuentes de información para los cálculos del IGAE tiene su base en agregaciones de la información de estadística básica, que proporcionan las encuestas sectoriales que genera el Instituto Nacional de Estadística, Geografía e Informática (INEGI), las diversas Cámaras Industriales y los registros administrativos de empresas públicas y privadas.

El IGAE también toma información de las encuestas y otras estadísticas del INEGI como son: la estadística de la Industria Minerometalúrgica, la Encuesta Industrial Mensual, la Estadística de la Industria Maquiladora de Exportación, la Encuesta de la Industria de la Construcción, la Encuesta Mensual sobre Establecimientos Comerciales y la Encuesta de Servicios.

Las fuentes de información proceden de Cámaras Industriales y de empresas, organismos privados como: Compañías Telefónicas, Industriales y de Aviación, Instituciones

Financieras y de Seguros, Empresas de Servicios Conexos al Transporte y Productoras de Automóviles y Camiones. Teléfonos de México, S.A. (TELMEX); Radio Móvil Dipsa, S.A. de C.V. (TELCEL); AVANTEL, S.A.; Grupo IUSACELL S.A. de C.V.; Aeropuertos y Servicios Auxiliares (ASA); Servicios a la Navegación en el Espacio Aéreo Mexicano (SENEAM); Compañía Mexicana de Aviación, S.A. de C.V. (MEXICANA); Aerovías de México, S.A. de C.V. (AEROMEXICO); Transportes Aeromar, S.A. de C.V.; Servicios Aéreos Litoral, S.A. de C.V.; Almacenadora Sur, S.A. de C.V.; Almacenadora Centro, S.A.; Servicios de Almacenamiento del Norte, S.A.; Asociación Mexicana de la Industria Automotriz, A.C. (AMIA); Asociación Nacional de Productores de Autobuses, Camiones y Tractocamiones, A.C. (ANPACT); Cementos Mexicanos (CEMEX); y Asociación Mexicana de Instituciones de Seguros (AMIS).

Se hace uso de los registros administrativos provenientes de las Instituciones Públicas entre las que se puede citar a: Petróleos Mexicanos (PEMEX); Comisión Federal de Electricidad (CFE); Secretaría de Comunicaciones y Transportes (SCT); Sistema de Transporte Colectivo Metro; Caminos y Puentes Federales (CAPUFE); Comisión Nacional del Agua (CONAGUA); Gobierno del Distrito Federal; Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE); Instituto Mexicano del Seguro Social (IMSS); Banco de México (BANXICO); Servicio Postal Mexicano (SEPOMEX); Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA); Secretaría de Gobernación (SG); Secretaría de Hacienda y Crédito Público (SHCP); Secretaría de Salud (SSA); y la Secretaría de Turismo (SECTUR).

De esta forma, los resultados del IGAE presentan una alta correlación con los del PIB Trimestral, tanto por su método de cálculo como por su elevada cobertura, la cual se ubica en alrededor de 84% del PIB Trimestral en el año de la base (1993), donde los sectores que se integran representan también por arriba del 80% cada uno.

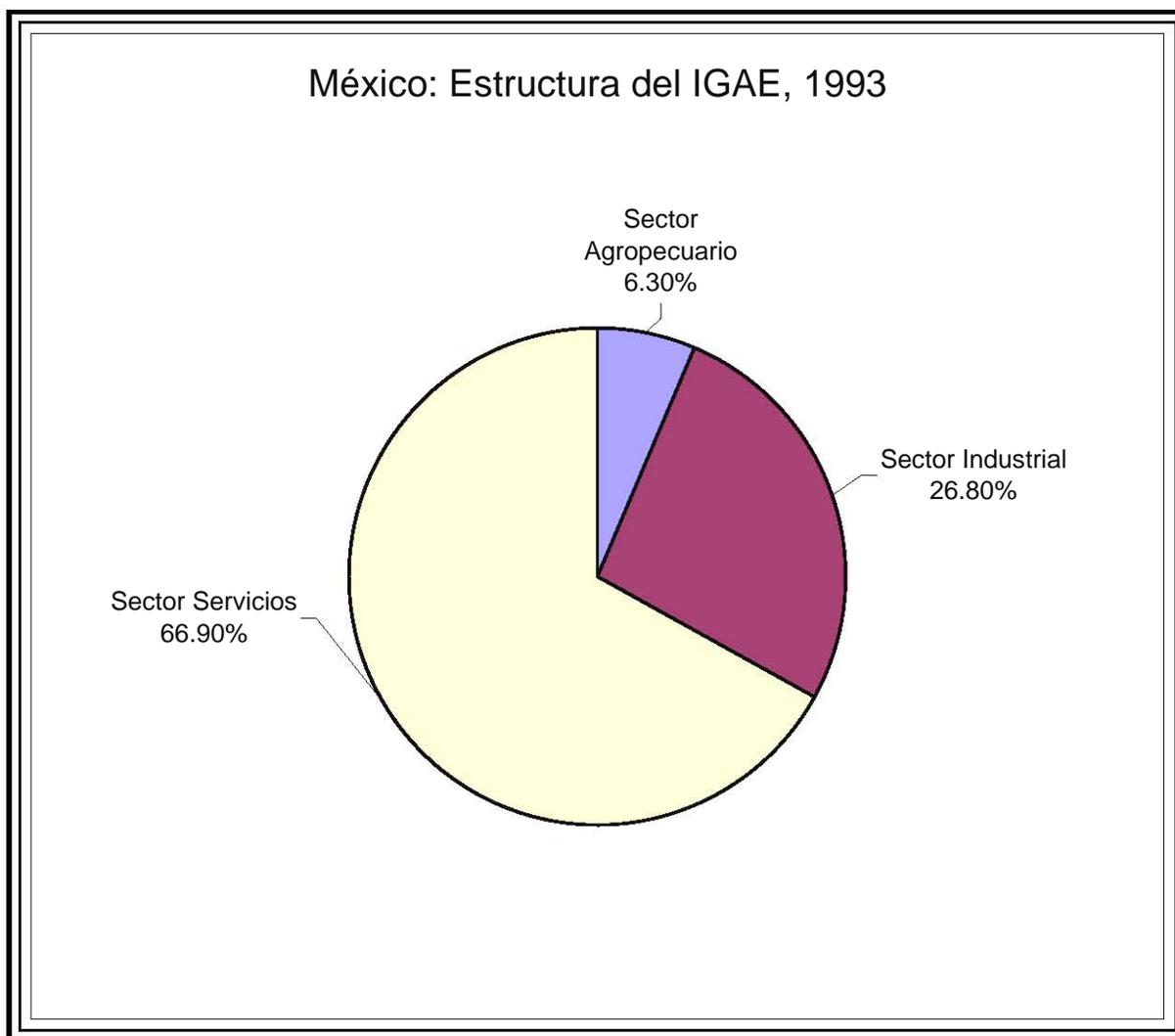
## **6.1. OBJETIVO DEL IGAE**

*El objetivo del IGAE es el de proporcionar información mensual y oportuna que permita un conocimiento sobre la evolución de la actividad económica del país.*

## 6.2. COMPONENTES DEL IGAE

El IGAE se estructura de la siguiente manera de acuerdo con su base que es 1993<sup>18</sup>: el Sector Agropecuario tiene una ponderación de 6.30%, el Sector Industrial de 26.80% y el Sector Servicios, que incluye a Comercio y Hoteles, pondera 66.90 por ciento.

**Ilustración 6.2. México: Estructura Porcentual del IGAE.**



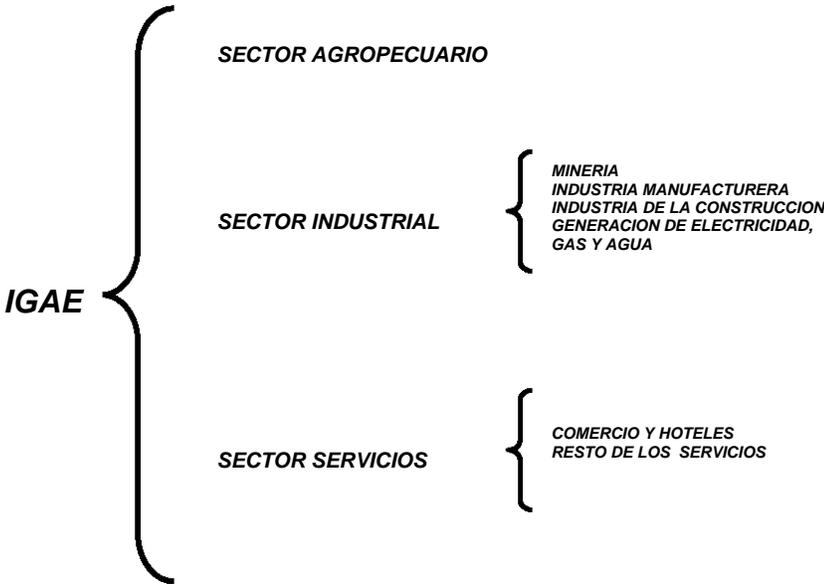
Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

El Indicador Global de la Actividad Económica contabiliza de manera específica al Sector Agropecuario, al Industrial y al de Servicios. Dentro del Industrial se contempla a la

<sup>18</sup> <http://dgcnesyp.inegi.gob.mx/cgi-win/bdieintsi.exe/MTDA050010>.

Minería, la Industria Manufacturera, la Industria de la Construcción, y la Generación de Electricidad, Gas y Agua. Dentro del Sector Servicios se calcula al Comercio, los Hoteles y el Resto de los Servicios.

**Ilustración 6.3. México: Sectores y Divisiones que Componen el IGAE.**

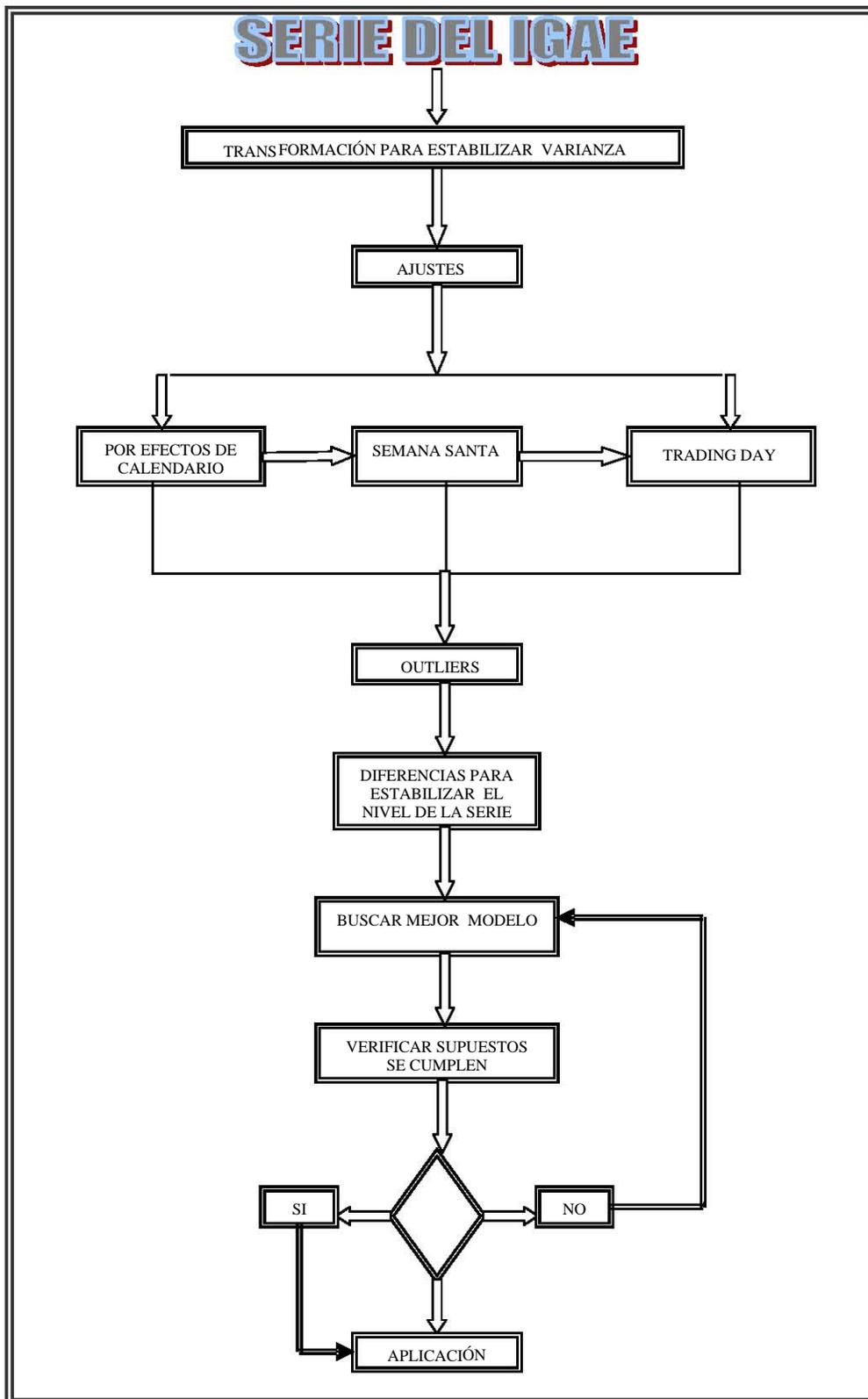


Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En la ilustración 6.3, observamos como se compone el Indicador Global de la Actividad Económica de forma desagregada.

Después de exponer las principales características del IGAE en términos generales. El diagrama siguiente muestra los pasos que haremos en el proceso de ajuste estacional de la serie real.

Ilustración 6.4. México: Diagrama de ajuste del IGAE.



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

## 7. AJUSTES PREVIOS PARA LA DESESTACIONALIZACIÓN

### 7.1. IGAE (TOTAL)

*El objetivo de este capítulo es aplicar la metodología expuesta en los apartados anteriores para ajustar la serie del Indicador Global de la Actividad Económica.*

En este apartado se analiza la serie para determinar por qué método (multiplicativo o aditivo) desestacionalizar la serie, se presenta la tabla 7.1 para seleccionar la potencia más adecuada para estabilizar la varianza de la serie, se selecciona la potencia que tenga el menor coeficiente de variación. Además se estiman si es que hay los ajustes previos de calendario (Trading Day y Easter) que se entienden como la frecuencia de los días de la semana y la presencia de la Semana Santa. Los outliers también son estimados y se definen como valores atípicos que pueden estar presentes en la serie, esto se realiza en el programa reg-ARIMA en X12-ARIMA.

Tabla 7.1. México: Cálculo para seleccionar la potencia de la transformación estabilizadora de la varianza del IGAE, 2000-2004.

*(Índice 1993 = 100)*

<b>INDICADOR GLOBAL DE LA ACTIVIDAD ECONOMICA (IGAE)</b>					
CALCULOS PARA SELECCIONAR LA POTENCIA DE LA TRANSFORMACION ESTABILIZADORA DE LA VARIANZA					
$(\lambda)$					
GRUPO	-1	-0.5	0	0.5	1
2000	0.0002	0.0028	0.0319	0.3607	4.0765
2001	0.0002	0.0024	0.0266	0.3001	3.3891
2002	0.0003	0.0029	0.0330	0.3742	4.2436
2003	0.0002	0.0029	0.0326	0.3720	4.2482
2004	0.0003	0.0031	0.0357	0.4163	4.8570
Coefficiente de Variación 2000-2004	0.0341	0.0492	0.0644	0.0795	0.0946

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

A continuación se presenta la serie del Indicador Global de la Actividad Económica con datos expresados en índices base 1993 con periodicidad mensual de 2000-2004, con sus respectivas tasas de crecimiento así como la gráfica de la serie.

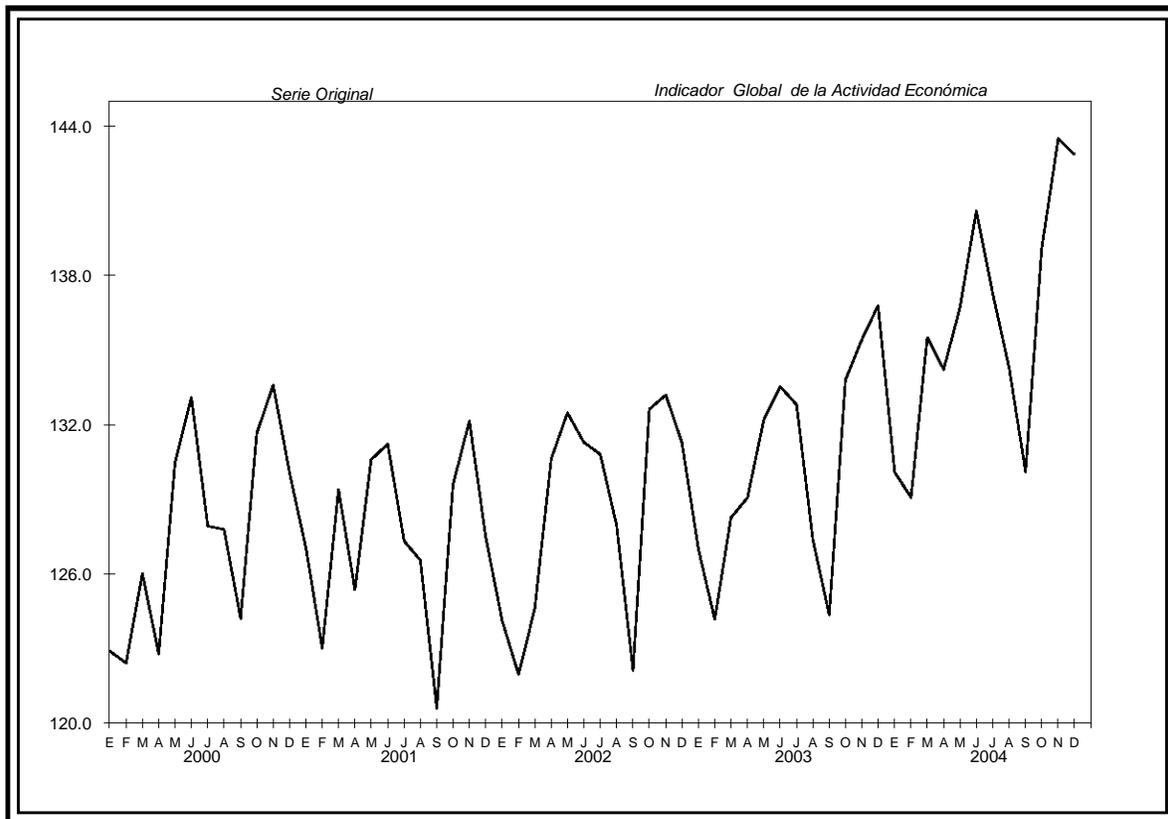
**Tabla 7.2. México: Serie Original del IGAE, 2000-2004.**  
(Índice 1993 = 100)

MÉXICO: INDICADOR GLOBAL DE LA ACTIVIDAD ECONÓMICA, 2000-2004. (1993=100)														
PERIODO	SERIE ORIGINAL		PERIODO	SERIE ORIGINAL		PERIODO	SERIE ORIGINAL		PERIODO	SERIE ORIGINAL		PERIODO	SERIE ORIGINAL	
	INDICE	VAR. ANUAL		INDICE	VAR. ANUAL		INDICE	VAR. ANUAL		INDICE	VAR. ANUAL		INDICE	VAR. ANUAL
2000 E	122.9		2001 E	127.1	3.37	2002 E	124.1	-2.32	2003 E	127.0	2.34	2004 E	130.1	2.43
F	122.4		F	123.0	0.49	F	122.0	-0.86	F	124.2	1.81	F	129.1	3.95
M	126.0		M	129.4	2.68	M	124.6	-3.69	M	128.3	2.92	M	135.5	5.65
A	122.8		A	125.4	2.10	A	130.6	4.19	A	129.1	-1.19	A	134.2	3.99
M	130.5		M	130.6	0.10	M	132.5	1.44	M	132.2	-0.21	M	136.7	3.40
J	133.1		J	131.2	-1.40	J	131.3	0.05	J	133.5	1.71	J	140.6	5.28
J	127.9		J	127.3	-0.47	J	130.8	2.74	J	132.8	1.52	J	137.2	3.35
A	127.8		A	126.6	-0.97	A	128.0	1.12	A	127.4	-0.46	A	134.3	5.42
S	124.2		S	120.6	-2.91	S	122.1	1.26	S	124.4	1.83	S	130.1	4.62
O	131.7		O	129.6	-1.58	O	132.6	2.34	O	133.8	0.89	O	139.1	3.96
N	133.6		N	132.1	-1.08	N	133.2	0.79	N	135.4	1.68	N	143.5	5.96
D	130.1		D	127.5	-1.97	D	131.3	2.96	D	136.8	4.19	D	142.9	4.44

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

### Ilustración 7.1. México: IGAE serie original, 2000-2004.

(Índice 1993 = 100)

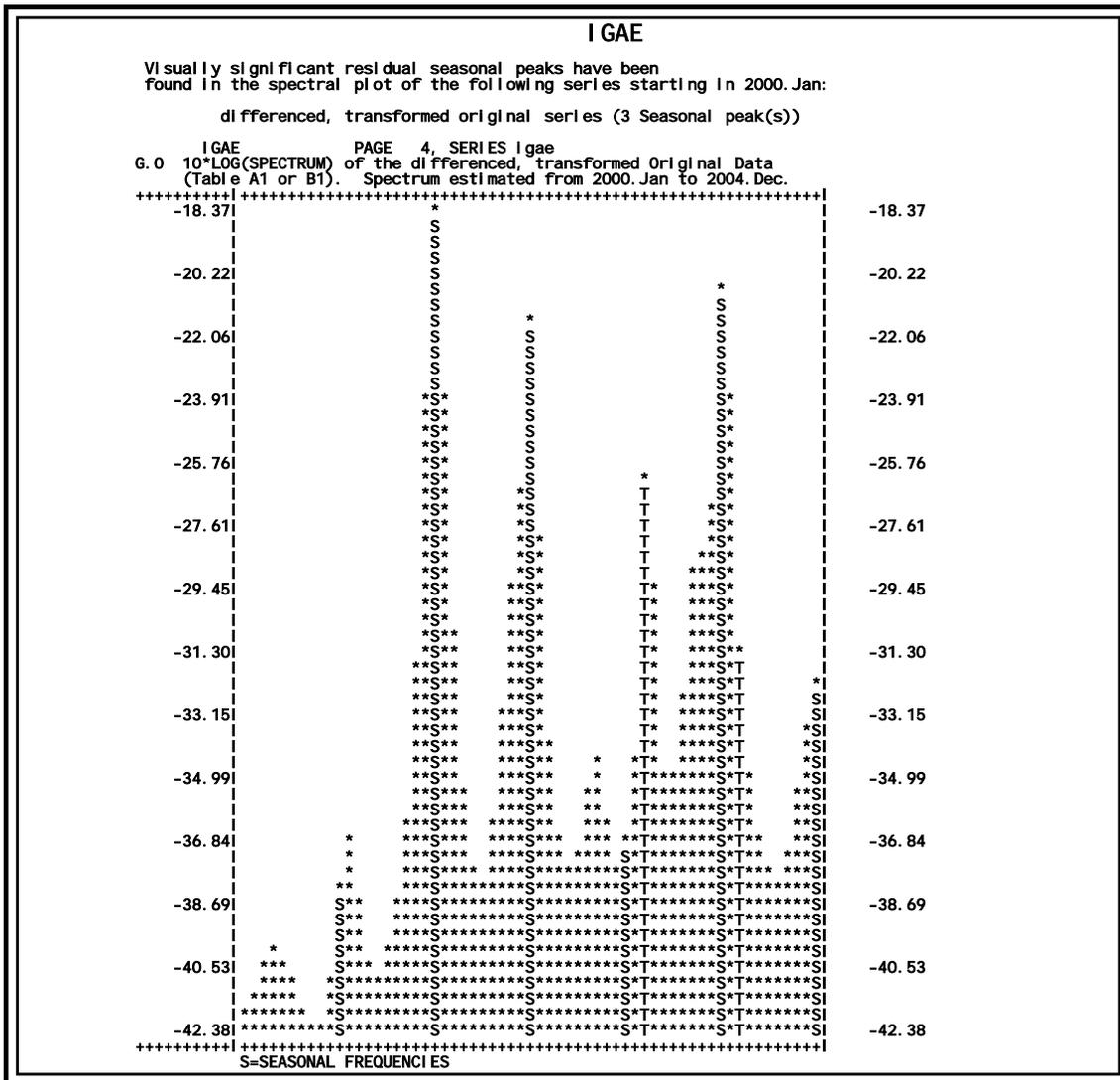


Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En un primer análisis se puede observar en la ilustración 7.1 que la serie es estacional en nivel porque tiene una tendencia lineal por lo que es necesario aplicar diferencias. Utilizaremos los datos transformados por logaritmos con la intención de estabilizar la varianza de la serie.

Se analiza el espectro de la serie original transformada por logaritmos la cual es tomada de los residuales y de la salida que el X12-ARIMA efectúa, el espectro es una representación gráfica de los datos transformados por logaritmos y se observa un comportamiento estacional porque presenta la serie tres picos estacionales y un pico de Trading Day. La gráfica muestra las variaciones anuales donde se puede observar que la serie presenta estacionalidad en el periodo en estudio. La gráfica y el espectro, justifican la necesidad de desestacionalizar la serie del IGAE en primera instancia.

Tabla 7.2. México: Espectro de la serie original transformada, 2000-2004.  
(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

## 7.2. EFECTOS DE CALENDARIO: SEMANA SANTA, DÍAS LABORABLES Y OUTLIERS

En este apartado determinaremos si la serie requiere que se ajuste por frecuencia de los días de la semana (td) y Semana Santa (easter) o si requiere algún ajuste por outliers, haremos uso del módulo reg-ARIMA, para estimarlos y ver si son significativos, que para el caso de Semana Santa se busca el número de días más significativo de tal efecto, es decir el número de días que la serie modifica su comportamiento previo al domingo de resurrección.

En la estimación de los efectos mencionados se utilizó la serie que se transformó por logaritmos de acuerdo con la transformación potencia más adecuada para la serie ver ilustración 7.1 donde se eligió la potencia más adecuada y puesto que el programa lo ejecuta automáticamente y lo representa como transformación logarítmica, la tabla siguiente describe los efectos.

**Tabla 7.3. México: Td y Easter, 2000-2004.**  
(Índice 1993 = 100)

```

INDICADOR GLOBAL DE LA ACTIVIDAD ECONOMICA

prefers log transformation

Irregular Regression Model
Trading Day + Easter[10]

Irregular Component Regression Span
From 2000.Jan to 2004.Dec

OUTLIER DETECTION
From 2000.Jan to 2004.Dec
Observations      60
Types              AO only
Method             add one
Critical |t| for AO outliers      3.69

Regression Model
-----
Variable           Parameter      Standard      t-value
                   Estimate      Error
-----
Trading Day
Mon                0.0580        0.04100       1.41
Tue                0.1187        0.03972       2.99
Wed                0.0020        0.03901       0.05
Thu                0.0466        0.04014       1.16
Fri                0.1110        0.03909       2.84
Sat                -0.0925       0.03865       -2.39
*Sun (derived)    -0.2437       0.04128       -5.90

Easter[10]         -0.6330       0.08371       -7.56

Automatically Identified Outliers
AO2000.Feb         -0.5441       0.11358       -4.79
AO2004.Feb         -0.5157       0.11413       -4.52
-----
*For full trading-day effects, the derived parameter estimate
is obtained indirectly as minus the sum of the directly estimated
parameters that define the effect.

Chi-squared Tests for Groups of Regressors
-----
Regression Effect      df      Chi-Square      P-Value
-----
Trading Day            6        158.46         0.00
-----

```

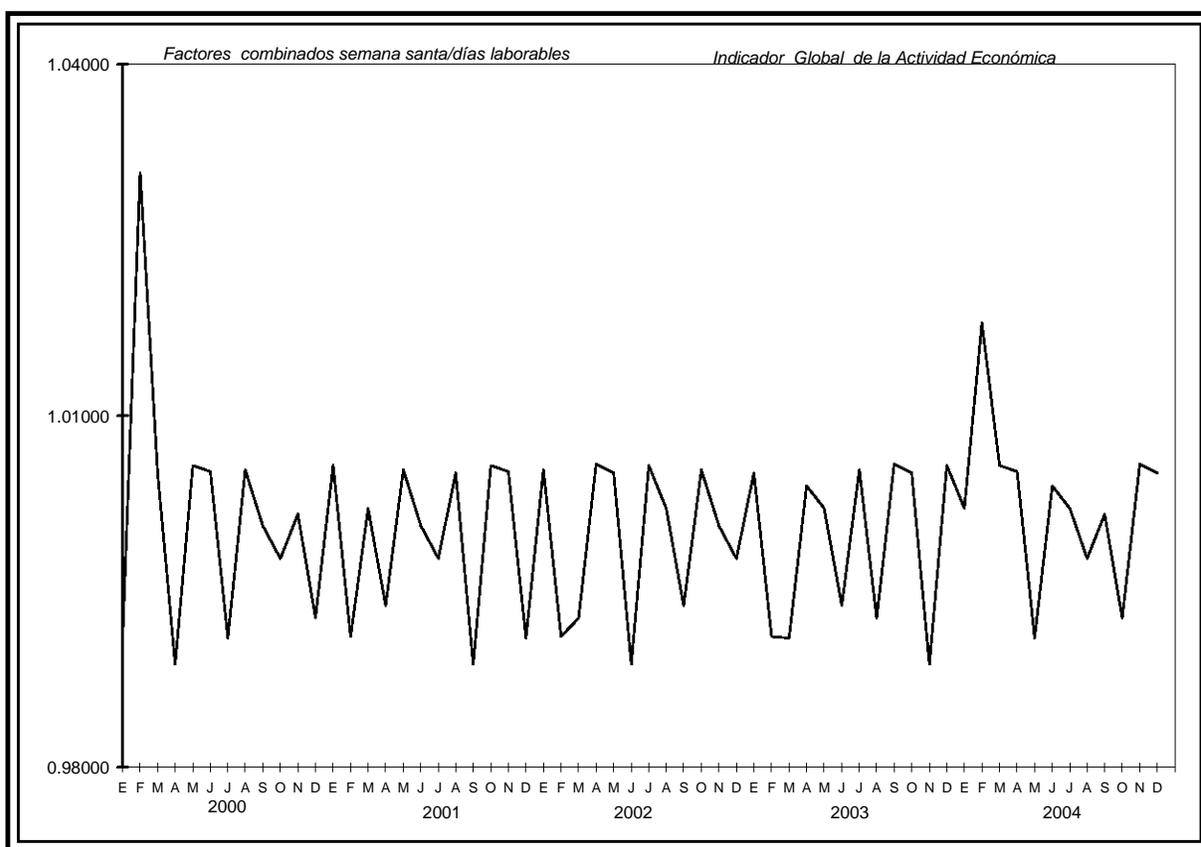
Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

La conclusión que se observa de la tabla 7.3 en los efectos de Semana Santa (Easter) y la frecuencia de los días de la semana (Trading Day) en el caso del trading day en la tabla podemos observar que los parámetros estimados para los días laborables el martes, viernes, sábado y el domingo son los de mayor significado para la serie en estudio. En el easter, el

número de días antes del domingo de pascua es 10, esto indica que la actividad del IGAE empieza a modificarse 10 días antes del domingo de pascua. Estos efectos tienen consecuencias para el mes en que caiga la semana santa (marzo o abril).

Con estos efectos que se estimaron, se procederá a ajustar la serie precisamente por *easter* y *td*, dividiéndose entre el producto de estos factores para tener la serie corregida. Para observar estos efectos presentamos la ilustración 7.2 de la estimación conjunta de estos dos factores (semana santa y días laborables).

**Ilustración 7.2. México: factores combinados semana santa/días laborables 2000-2004.**  
(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

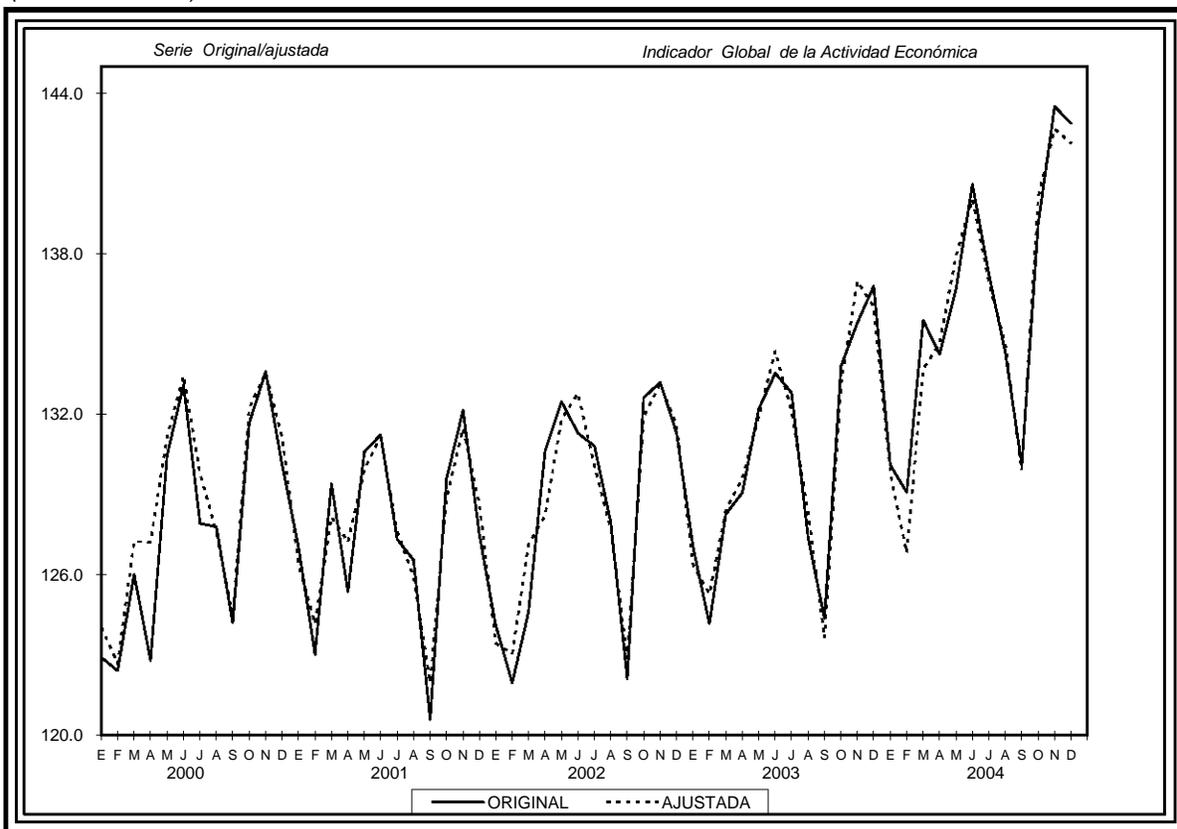
Debido a que se realizó el ajuste por los factores de Semana Santa y días laborables en la tabla 7.4 presentamos la serie corregida por estos factores; así como también la ilustración 7.3 para tener otra apreciación de la magnitud del cambio que se efectuó.

Tabla 7.4. México: Serie original corregida por días laborables y semana santa, 2000-2004.  
(Índice 1993 = 100)

MEXICO: INDICADOR GLOBAL DE LA ACTIVIDAD ECONÓMICA 2000-2004. (1993=100)														
PERIODO	SERIE ORIGINAL CORREGIDA POR D/SS		PERIODO	SERIE ORIGINAL CORREGIDA POR D/SS		PERIODO	SERIE ORIGINAL CORREGIDA POR D/SS		PERIODO	SERIE ORIGINAL CORREGIDA POR D/SS		PERIODO	SERIE ORIGINAL CORREGIDA POR D/SS	
	INDICE	VAR. ANUAL		INDICE	VAR. ANUAL		INDICE	VAR. ANUAL		INDICE	VAR. ANUAL		INDICE	VAR. ANUAL
2000 E	124.0		2001 E	126.3	1.85	2002 E	123.4	-2.28	2003 E	126.4	2.37	2004 E	129.8	2.74
F	118.8		F	123.2	3.74	F	123.0	-0.13	F	125.3	1.81	F	126.8	1.22
M	124.4		M	128.1	2.99	M	127.1	-0.77	M	128.4	0.99	M	133.7	4.12
A	125.2		A	127.2	1.58	A	128.2	0.80	A	129.6	1.09	A	134.6	3.86
M	129.7		M	129.9	0.14	M	131.8	1.46	M	131.9	0.10	M	137.9	4.56
J	132.4		J	131.1	-0.95	J	132.8	1.25	J	134.4	1.20	J	140.0	4.21
J	129.1		J	127.6	-1.14	J	130.1	1.93	J	132.1	1.56	J	137.0	3.69
A	127.1		A	125.9	-0.95	A	127.7	1.43	A	128.3	0.48	A	134.6	4.88
S	124.1		S	122.0	-1.75	S	122.9	0.75	S	123.6	0.61	S	129.9	5.07
O	132.0		O	128.8	-2.35	O	131.9	2.37	O	133.1	0.92	O	140.1	5.26
N	133.4		N	131.5	-1.44	N	133.1	1.25	N	137.0	2.89	N	142.7	4.16
D	131.0		D	128.7	-1.80	D	131.6	2.26	D	136.0	3.37	D	142.1	4.51

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Ilustración 7.3. México: Serie original y la serie ajustada por semana santa y efectos de calendario 2000-2004.  
(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Habiendo analizado la serie por (efectos de calendario) procederemos a detectar si la serie presenta outliers que se entienden como valores atípicos dentro de la serie y que pueden

ser cambio de nivel, cambio temporal y aditivos, para lo cual utilizaremos el reg-ARIMA en X12-ARIMA, para esto utilizamos la serie que ajustamos por días laborables y semana santa, el resultado lo presentamos en la siguiente tabla.

**Tabla 7.5. México: Identificación de outliers, 2000-2004.**  
(Índice 1993 = 100)

```

Irregular Regression Model
Trading Day + Easter[10]

Irregular Component Regression Span
From 2000. Jan to 2004. Dec

OUTLIER DETECTION
From 2000. Jan to 2004. Dec
Observations      60
Types             AO only
Method            add one
Critical |t| for AO outliers      3.69

-----
OUTLIER DETECTION
From 2000. Jan to 2004. Dec
Observations      60
Types             AO, LS and TC
Method            add one
Critical |t| for AO outliers      3.69
Critical |t| for LS outliers      3.69
Critical |t| for TC outliers      3.69
Regression Model

-----
Variable          Parameter      Standard
                  Estimate      Error      t -value
-----
Automatically Identified Outliers
TC2000. Feb      -0.0331      0.00755      -4.39
    
```

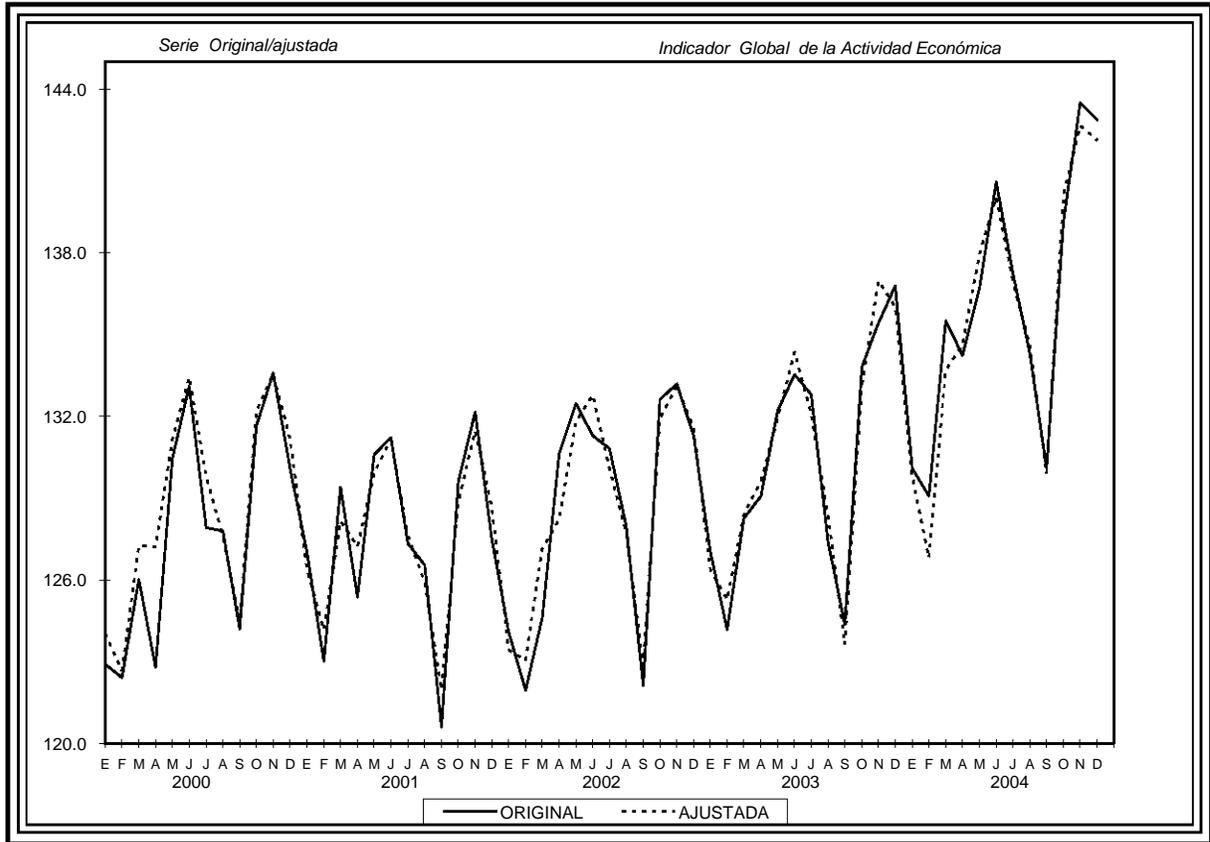
Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Los resultados obtenidos en la tabla 7.5 demuestran la necesidad de incluir un outliers denominado como *cambio temporal* para corregir la serie en ese período, TC2000 febrero, es decir una variable que corrija el cambio temporal y que ocurrió para ese período.

Al incluir Semana Santa, días laborables (*td*) y outliers presentamos en la ilustración 7.4 a las dos series: la original y la ajustada por los efectos antes mencionados.

### Ilustración 7.4. México: Serie original y la serie corregida por semana santa, Trading day y outliers, 2000-2005.

(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Con la estimación de los distintos efectos que contiene la serie como son la Semana Santa, los días laborables y los outliers ésta serie ajustada ya podrá ser representada por un proceso ARIMA adecuado.

### 7.3. IDENTIFICACIÓN DE LA MEJOR DIFERENCIA PARA LA SERIE AJUSTADA DEL IGAE

En este apartado para hallar el mejor modelo se recurre a la metodología tratada en el punto cinco y utilizaremos la serie ajustada por los efectos de Semana Santa, días laborables y outliers (observaciones atípicas), así en primera instancia buscaremos la mejor diferencia para volver estacionaria la serie, de esta forma se presentan las tablas de las distintas diferencias que se aplicaron a la serie.

En la práctica por lo general no son necesarias diferencias mayores a dos para hacer estacionarias a las series y en nuestro caso no fue la excepción ya que utilizamos diferencias *ordinarias* y *estacionales* de un grado, esto lo constatamos en las tablas 7.6, 7.7 y 7.8 respectivamente.

Tabla 7.6. México: Diferencia estacional (0 1) de la serie transformada y ajustada por semana santa, trading day y outliers, 2000-2004.

(Índice 1993 = 100)

MODEL IDENTIFICATION (Using regression residuals)												
Differencing: Seasonal Order=1												
Sample Autocorrelations of the Residuals												
Lag	1	2	3	4	5	6	7	8	9	10	11	12
ACF	0.88	0.81	0.72	0.63	0.54	0.44	0.30	0.22	0.14	0.05	0.02	-0.05
SE	0.14	0.23	0.28	0.32	0.34	0.36	0.37	0.38	0.38	0.38	0.38	0.38
Q	39.95	73.95	101.31	122.84	139.30	150.18	155.55	158.35	159.60	159.73	159.76	159.91
DF	1	2	3	4	5	6	7	8	9	10	11	12
P	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Lag	13	14	15	16	17	18	19	20	21	22	23	24
ACF	-0.07	-0.06	-0.04	-0.02	0.02	0.02	0.04	0.06	0.05	0.04	0.02	-0.03
SE	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38
Q	160.20	160.48	160.62	160.66	160.68	160.71	160.81	161.10	161.29	161.46	161.50	161.61
DF	13	14	15	16	17	18	19	20	21	22	23	24
P	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

The P-values approximate the probability of observing a Q-value at least this large when the model fitted is correct. When DF is positive, small values of P, customarily those below 0.05, indicate model inadequacy.

Sample Autocorrelations of the Residuals												
	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0	
1												0.885
2												0.807
3												0.716
4												0.628
5												0.543
6												0.436
7												0.303
8												0.216
9												0.142
10												0.045
11												0.021
12							X					-0.048
13							XX					-0.065
14							XX					-0.063
15							X					-0.043
16							X					-0.023
17												0.019
18												0.016
19							X					0.035
20							X					0.059
21							X					0.046
22							X					0.044
23												0.018
24							X					-0.033

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Tabla 7.7. México: Diferencia no estacional (1 0) de la serie transformada y ajustada por semana santa, trading day y outliers, 2000-2004.

(Índice 1993 = 100)

Differencing: Nonseasonal Order=1  
Sample Autocorrelations of the Residuals

Lag	1	2	3	4	5	6	7	8	9	10	11	12
ACF	0.03	-0.32	-0.47	-0.10	0.36	0.06	0.35	-0.07	-0.42	-0.29	0.03	0.74
SE	0.13	0.13	0.14	0.17	0.17	0.18	0.18	0.19	0.19	0.21	0.22	0.22
Q	0.05	6.61	20.88	21.47	30.06	30.34	38.96	39.35	52.25	58.57	58.63	101.04
DF	1	2	3	4	5	6	7	8	9	10	11	12
P	0.820	0.037	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Lag	13	14	15	16	17	18	19	20	21	22	23	24
ACF	0.04	-0.24	-0.38	-0.08	0.27	0.04	0.26	-0.04	-0.30	-0.23	0.00	0.55
SE	0.26	0.26	0.26	0.27	0.27	0.27	0.27	0.28	0.28	0.28	0.29	0.29
Q	101.16	105.80	117.86	118.41	124.51	124.68	130.80	130.96	139.77	144.88	144.89	176.43
DF	13	14	15	16	17	18	19	20	21	22	23	24
P	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

The P-values approximate the probability of observing a Q-value at least this large when the model fitted is correct. When DF is positive, small values of P, customarily those below 0.05, indicate model inadequacy.

Sample Autocorrelations of the Residuals

	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
1						X					0.029
2					XXXXXXX						-0.322
3				XXXXXXXXXX							-0.471
4					XX						-0.095
5						XXXXXXXXXX					0.359
6						XX					0.065
7						XXXXXXXXXX					0.353
8						XX					-0.074
9				XXXXXXXXXX							-0.423
10				XXXXXXXXXX							-0.293
11						X					0.029
12						XXXXXXXXXXXXXXXXXXXX					0.744
13						X					0.039
14					XXXXXX						-0.241
15				XXXXXXXXXX							-0.384
16					XX						-0.081
17						XXXXXX					0.267
18						X					0.044
19						XXXXXX					0.261
20						X					-0.043
21				XXXXXXXXXX							-0.305
22				XXXXXX							-0.229
23											-0.003
24						XXXXXXXXXXXX					0.554

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Tabla 7.8. México: Diferencia no estacional y estacional (1 1) de la serie transformada y ajustada por semana santa, trading day y outliers, 2000-2004.

(Índice 1993 = 100)

Differencing: Nonseasonal Order=1, Seasonal Order=1												
Sample Autocorrelations of the Residuals												
Lag	1	2	3	4	5	6	7	8	9	10	11	12
ACF	-0.18	-0.01	0.00	0.01	0.21	0.10	-0.29	-0.02	0.11	-0.19	0.12	-0.33
SE	0.15	0.15	0.15	0.15	0.15	0.16	0.16	0.17	0.17	0.17	0.18	0.18
Q	1.57	1.58	1.58	1.58	4.05	4.58	9.53	9.56	10.27	12.57	13.54	20.89
DF	1	2	3	4	5	6	7	8	9	10	11	12
P	0.210	0.455	0.665	0.813	0.542	0.599	0.216	0.298	0.329	0.249	0.259	0.052

Lag	13	14	15	16	17	18	19	20	21	22	23	24
ACF	-0.10	0.06	0.00	-0.12	0.16	-0.10	0.07	0.13	-0.08	0.08	0.08	0.03
SE	0.19	0.19	0.19	0.19	0.19	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Q	21.51	21.74	21.75	22.77	24.63	25.37	25.80	27.17	27.68	28.22	28.86	28.95
DF	13	14	15	16	17	18	19	20	21	22	23	24
P	0.063	0.084	0.115	0.120	0.103	0.115	0.136	0.130	0.149	0.169	0.185	0.222

The P-values approximate the probability of observing a Q-value at least this large when the model fitted is correct. When DF is positive, small values of P, customarily those below 0.05, indicate model inadequacy.

Sample Autocorrelations of the Residuals

	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
1											-0.177
2											-0.008
3											0.002
4											0.005
5											0.212
6											0.096
7											-0.294
8											-0.019
9											0.109
10											-0.192
11											0.123
12											-0.334
13											-0.096
14											0.058
15											-0.004
16											-0.117
17											0.156
18											-0.097
19											0.072
20											0.127
21											-0.076
22											0.076
23											0.082
24											0.030

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En conclusión de las tablas anteriores podemos decir que la diferencia (0 1) que corresponde a una diferencia estacional no cumple con el objetivo de volver estacionaria la serie por la razón de que tiene una desviación estándar muy alta, en relación con la diferencia (1 0), que corresponde a una diferencia no estacional y (1 1), *muestra que necesitamos una diferencia no estacional y una diferencia estacional.*

Utilizaremos la tabla 7.8 que contiene la diferencia (1 1) donde observamos que existe una fuerte correlación entre los datos y tiene la menor desviación estándar condición para considerarla como apropiada, sin embargo en los retrasos 7 y 12 posiblemente se requiera incluirlos en el modelo, al 7 como promedio móvil *ordinarios* y en el retraso 12 podemos probar introduciendo un parámetro de promedio móvil *estacional*. Con estos argumentos podemos proponer un posible modelo ARIMA que ajuste correctamente a la serie del IGAE, el modelo propuesto es (0 1 0) (0 1 1).

## 8. BÚSQUEDA DEL MEJOR MODELO

Con los resultados anteriores y con la primer propuesta de modelo ARIMA podemos estimar el modelo (0 1 0) (0 1 1) y verificar si cumple con los supuestos de Box y Jenkins, que pueden ser consultados en el anexo de este trabajo, si no cumple buscar otro modelo hasta encontrar el más adecuado. Veamos los resultados de la siguiente tabla.

Tabla 8.1. México: Estimación de los parámetros para el modelo (0 1 0) (0 1 1), 2000-2004.  
(Índice 1993 = 100)

```

T GAE
C 15 Final irregular component regression
Irregular Regression Model
Trading Day + Easter[10]
Irregular Component Regression Span
From 2000.Jan to 2004.Dec

OUTLIER DETECTION
From 2000.Jan to 2004.Dec
Observations 60
Types AO only
Method add one
Critical |t| for AO outliers 3.69

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
Trading Day
Mon 0.0580 0.04100 1.41
Tue 0.1187 0.03972 2.99
Wed 0.0020 0.03901 0.05
Thu 0.0466 0.04014 1.16
Fri 0.1110 0.03909 2.84
Sat -0.0925 0.03865 -2.39
*Sun (derived) -0.2437 0.04128 -5.90
Easter[10] -0.6330 0.08371 -7.56

Automatically Identified Outliers
A02000.Feb -0.5441 0.11358 -4.79
A02004.Feb -0.5157 0.11413 -4.52

*For full trading-day effects, the derived parameter estimate
is obtained indirectly as minus the sum of the directly estimated
parameters that define the effect.

Chi-squared Tests for Groups of Regressors
-----
Regression Effect df Chi-Square P-Value
-----
Trading Day 6 158.46 0.00

Variance 0.11322E-01
MODEL DEFINITION
Transformation
Log(y)

Regression Model
TC2000.02

ARIMA Model
(0 1 0)(0 1 1)

regARIMA Model Span
From 2000.Jan to 2004.Dec

MODEL ESTIMATION/EVALUATION
Exact ARIMA likelihood estimation
Max total ARMA iterations 200
Max ARMA iter's w/in an IGLS iterati 40
Convergence tolerance 1.00E-05

Average absolute percentage error in within-sample forecasts:
Last year: 1.08 Last-1 year: 0.73 Last-2 year: 1.64
Last three years: 1.15

Estimation converged in 4 ARMA iterations, 19 function evaluations.

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
TC2000.02 -0.0326 0.00742 -4.40

ARIMA Model: (0 1 0)(0 1 1)
Nonseasonal differences: 1
Seasonal differences: 1

Parameter Estimate Standard
Errors
-----
Seasonal MA
Lag 12 0.4911 0.14484

Variance 0.64746E-04

```

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En la tabla 8.1 estimamos un easter (10) que son los días en que la serie presenta una modificación antes del domingo de resurrección, también la serie presenta trading day, estos dos efectos no cambiaran cuando estimemos los demás parámetros, se observa que el parámetro MA(12) de promedio móvil estacional, es significativo debido a que el estadístico “t” que se obtiene al dividir el valor estimado entre el error estándar es mayor a 2.0 de acuerdo con la teoría estadística.

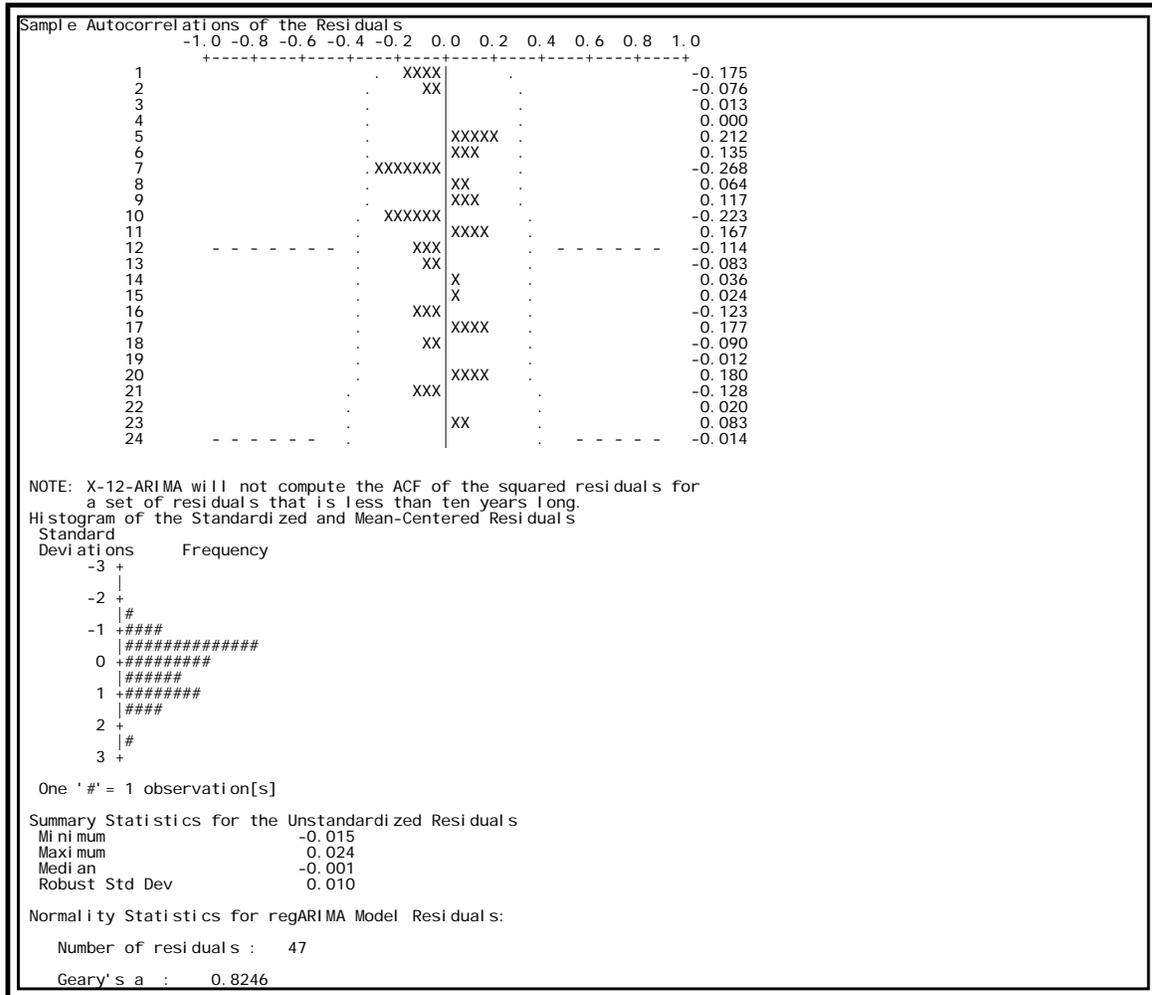
Tabla 8.2. México: Correlación de los parámetros para el modelo (0 1 0) (0 1 1), 2000-2004. (Índice 1993 = 100)

Likelihood Statistics												
Effective number of observations (nefobs)												47
Number of parameters estimated (np)												3
Log likelihood												158.3182
Transformation Adjustment												-229.0007
Adjusted Log likelihood (L)												-70.6825
AIC												147.3649*
AICC (F-corrected-AIC)												147.9231*
Hannan Quinn												149.4536*
BIC												152.9154*
* NOTE: These statistics do not contain a penalty for parameters estimated by x11regression to produce the prior adjustment factors because the x11regression estimates are not maximum likelihood estimates. Therefore they cannot be compared to the statistics from models in which regression variables in a regARIMA model are used to estimate the same effects.												
DIAGNOSTIC CHECKING												
Sample Autocorrelations of the Residuals												
Lag	1	2	3	4	5	6	7	8	9	10	11	12
ACF	-0.17	-0.08	0.01	0.00	0.21	0.13	-0.27	0.06	0.12	-0.22	0.17	-0.11
SE	0.15	0.15	0.15	0.15	0.15	0.16	0.16	0.17	0.17	0.17	0.18	0.18
Q	1.53	1.83	1.84	1.84	4.31	5.32	9.46	9.70	10.53	13.63	15.42	16.28
DF	0	1	2	3	4	5	6	7	8	9	10	11
P	0.000	0.176	0.399	0.607	0.366	0.378	0.149	0.206	0.230	0.136	0.117	0.131
Lag	13	14	15	16	17	18	19	20	21	22	23	24
ACF	-0.08	0.04	0.02	-0.12	0.18	-0.09	-0.01	0.18	-0.13	0.02	0.08	-0.01
SE	0.18	0.18	0.18	0.18	0.18	0.19	0.19	0.19	0.19	0.19	0.19	0.20
Q	16.75	16.83	16.88	18.00	20.40	21.04	21.05	23.80	25.24	25.28	25.94	25.96
DF	12	13	14	15	16	17	18	19	20	21	22	23
P	0.159	0.207	0.263	0.263	0.203	0.225	0.277	0.204	0.192	0.235	0.254	0.303
The P-values approximate the probability of observing a Q-value at least this large when the model fitted is correct. When DF is positive, small values of P, customarily those below 0.05, indicate model inadequacy.												

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En la tabla 8.2 observamos que no existe correlación en el parámetro estimado MA(12) estacional, en las autocorrelaciones de los residuales (ACF) se observa que no es mayor que la desviación estándar (SE).

Tabla 8.3. México: Autocorrelación de los residuales para el modelo (0 1 0) (0 1 1), 2000-2004.  
(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En la tabla 8.3 observamos que el modelo propuesto (0 1 0) (0 1 1) en el histograma presenta una distribución normal ya que su media y su desviación estándar tienden a cero. Además de que en las autocorrelaciones simples de los residuales ningún retraso se sale de las bandas.



Tabla 8.5. México: estimación del modelo (0 1 0) (1 1 0), 2000-2004.  
(Índice 1993 = 100)

```

IGAE
C 15 Final irregular component regression
Irregular Regression Model
Trading Day + Easter[10]

Irregular Component Regression Span
From 2000.Jan to 2004.Dec

OUTLIER DETECTION
From 2000.Jan to 2004.Dec
Observations 60
Types AO only
Method add one
Critical |t| for AO outliers 3.69

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
Trading Day
Mon 0.0580 0.04100 1.41
Tue 0.1187 0.03972 2.99
Wed 0.0020 0.03901 0.05
Thu 0.0466 0.04014 1.16
Fri 0.1110 0.03909 2.84
Sat -0.0925 0.03865 -2.39
*Sun (derived) -0.2437 0.04128 -5.90
Easter[10] -0.6330 0.08371 -7.56

Automatically Identified Outliers
A02000.Feb -0.5441 0.11358 -4.79
A02004.Feb -0.5157 0.11413 -4.52
-----
*For full trading-day effects, the derived parameter estimate
is obtained indirectly as minus the sum of the directly estimated
parameters that define the effect.

Chi-squared Tests for Groups of Regressors
-----
Regression Effect df Chi-Square P-Value
-----
Trading Day 6 158.46 0.00

Variance 0.11322E-01
MODEL DEFINITION
Transformation
Log(y)

Regression Model
TC2000.02

ARIMA Model
(0 1 0)(1 1 0)

regARIMA Model Span
From 2000.Jan to 2004.Dec

MODEL ESTIMATION/EVALUATION
Exact ARIMA likelihood estimation 200
Max total ARMA iterations 40
Max ARMA iter's w/in an IGLS iterati 40
Convergence tolerance 1.00E-05

NOTE: Insufficient data to compute average forecast error diagnostic.
Estimation converged in 10 ARMA iterations, 39 function evaluations.

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
TC2000.02 -0.0346 0.00748 -4.63

ARIMA Model: (0 1 0)(1 1 0)
Nonseasonal differences: 1
Seasonal differences: 1

Parameter Estimate Standard
Errors
-----
Seasonal AR
Lag 12 -0.4382 0.15779
Variance 0.66118E-04

Likelihood Statistics
-----
Effective number of observations (nefobs) 47
Number of parameters estimated (np) 3
Log likelihood 158.1963
Transformation Adjustment -229.0007
Adjusted Log likelihood (L) -70.8044
AIC 147.6087*
AICC (F-corrected-AIC) 148.1669*
Hannan Quinn 149.6974*
BIC 153.1592*
-----
* NOTE: These statistics do not contain a penalty for parameters
estimated by x11regression to produce the prior adjustment
factors because the x11regression estimates are not maximum
likelihood estimates. Therefore they cannot be compared to
the statistics from models in which regression variables in
a regARIMA model are used to estimate the same effects.

```

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).



Se descarta el modelo propuesto (0 1 0) (1 1 0) porque presenta en el espectro un pico estacional, el retraso 7 tiene una correlación muy alta lo consideraremos en la estimación de otro modelo con la finalidad de eliminar el pico estacional.

El modelo a estimar es (0 1 [7]) (1 1 0), a continuación presentamos las tablas 8.7, 8.8 y 8.9 respectivamente.

Tabla 8.7. México: estimación de los parámetros para el modelo (0 1 [7]) (1 1 0), 2000-2004. (Índice 1993 = 100)

```

C 15 Final Irregular component regression
Irregular Regression Model
Trading Day + Easter[10]
Irregular Component Regression Span
From 2000.Jan to 2004.Dec

OUTLIER DETECTION
From 2000.Jan to 2004.Dec
Observations 60
Types AO only
Method add one
Critical |t| for AO outliers 3.69

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
Trading Day
Mon 0.0580 0.04100 1.41
Tue 0.1187 0.03972 2.99
Wed 0.0020 0.03901 0.05
Thu 0.0466 0.04014 1.16
Fri 0.1110 0.03909 2.84
Sat -0.0925 0.03865 -2.39
*Sun (derived) -0.2437 0.04128 -5.90
Easter[10] -0.6330 0.08371 -7.56

Automatically Identified Outliers
A02000.Feb -0.5441 0.11358 -4.79
A02004.Feb -0.5157 0.11413 -4.52

*For full trading-day effects, the derived parameter estimate
is obtained indirectly as minus the sum of the directly estimated
parameters that define the effect.

Chi-squared Tests for Groups of Regressors
-----
Regression Effect df Chi-Square P-Value
-----
Trading Day 6 158.46 0.00

Variance 0.11322E-01
MODEL DEFINITION
Transformation
Log(y)

Regression Model
TC2000.02

ARIMA Model
(0 1 [7])(1 1 0)

regARIMA Model Span
From 2000.Jan to 2004.Dec

MODEL ESTIMATION/EVALUATION
Exact ARIMA likelihood estimation
Max total ARMA iterations 200
Max ARMA iter's w/in an IGLS iteration 40
Convergence tolerance 1.00E-05

NOTE: Insufficient data to compute average forecast error diagnostic.
Estimation converged in 10 ARMA iterations, 49 function evaluations.

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
TC2000.02 -0.0324 0.00720 -4.50

ARIMA Model: (0 1 [7])(1 1 0)
Nonseasonal differences: 1
Seasonal differences: 1

Parameter Estimate Standard
Errors
-----
Seasonal AR
Lag 12 -0.4239 0.15804

Nonseasonal MA
Lag 7 0.2972 0.13808

Variance 0.60523E-04

```

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

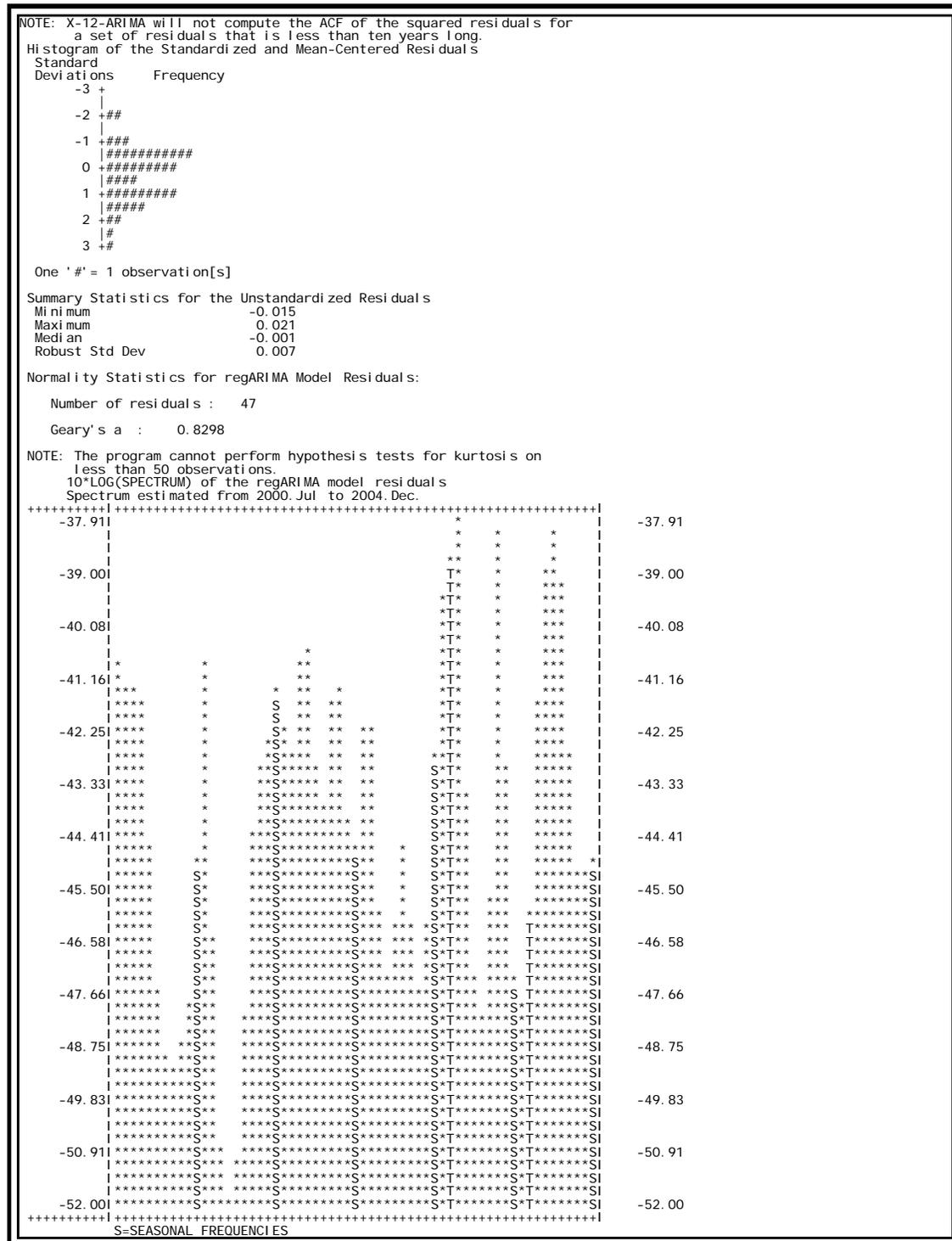
Tabla 8.8. México: parámetros estimados para el modelo (0 1 [7]) (1 1 0), 2000-2004.  
(Índice 1993 = 100)

ARMA Parameter Correlation matrix	
Parameter	1 2
-----	
Seasonal AR	
Lag 12	1.00
Nonseasonal MA	
Lag 7	0.10 1.00
-----	
Likelihood Statistics	
Effective number of observations (nefobs)	47
Number of parameters estimated (np)	4
Log Likelihood	160.0415
Transformation Adjustment	-229.0007
Adjusted Log Likelihood (L)	-68.9592
AIC	145.9184*
AICC (F-corrected-AIC)	146.8708*
Hannan Quinn	148.7033*
BIC	153.3190*
-----	
* NOTE: These statistics do not contain a penalty for parameters estimated by x11regression to produce the prior adjustment factors because the x11regression estimates are not maximum likelihood estimates. Therefore they cannot be compared to the statistics from models in which regression variables in a regARIMA model are used to estimate the same effects.	
-----	
DIAGNOSTIC CHECKING	
Sample Autocorrelations of the Residuals	
Lag	1 2 3 4 5 6 7 8 9 10 11 12
ACF	-0.24 -0.02 -0.07 -0.01 0.18 0.13 -0.12 -0.07 0.11 -0.20 0.16 -0.01
SE	0.15 0.15 0.15 0.15 0.15 0.16 0.16 0.16 0.16 0.17 0.17 0.17
Q	2.93 2.94 3.17 3.17 4.85 5.81 6.60 6.92 7.72 10.27 11.83 11.83
DF	0 0 1 2 3 4 5 6 7 8 9 10
P	0.000 0.000 0.075 0.205 0.183 0.214 0.252 0.328 0.358 0.247 0.223 0.296
Lag	13 14 15 16 17 18 19 20 21 22 23 24
ACF	-0.01 -0.07 0.10 -0.18 0.11 0.01 -0.08 0.22 -0.12 0.04 0.00 -0.05
SE	0.17 0.17 0.17 0.18 0.18 0.18 0.18 0.18 0.19 0.19 0.19 0.19
Q	11.84 12.17 12.93 15.39 16.40 16.40 16.89 21.01 22.26 22.39 22.39 22.64
DF	11 12 13 14 15 16 17 18 19 20 21 22
P	0.376 0.432 0.453 0.352 0.356 0.425 0.462 0.279 0.272 0.320 0.377 0.422
The P-values approximate the probability of observing a 0-value at least this large when the model fitted is correct. When DF is positive, small values of P, customarily those below 0.05, indicate model inadequacy.	
Sample Autocorrelations of the Residuals	
	-1.0 -0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
1	.XXXXXXXXX . -0.242
2	. . . -0.015
3	. . . XX . -0.066
4	. . . . . -0.005
5	. . . . . XXXX . 0.175
6	. . . . . XXX . 0.131
7	. . . . . XXX . -0.117
8	. . . . . XX . -0.073
9	. . . . . XXX . 0.115
10	. . . . . XXXXX . -0.202
11	. . . . . XXXX . 0.156
12	. . . . . . . . . -0.011
13	. . . . . . . . . -0.006
14	. . . . . . . . . -0.069
15	. . . . . . . . . 0.103
16	. . . . . . . . . -0.182
17	. . . . . . . . . 0.115
18	. . . . . . . . . 0.007
19	. . . . . . . . . -0.077
20	. . . . . . . . . 0.220
21	. . . . . . . . . -0.119
22	. . . . . . . . . 0.038
23	. . . . . . . . . 0.003
24	. . . . . . . . . -0.050

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Tabla 8.9. México: histograma y espectro del modelo (0 1 [7]) (1 1 0), 2000-2004.

(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Al estimar el modelo  $(0 \ 1 \ [7]) \ (1 \ 1 \ 0)$ , lo consideramos como óptimo por las siguientes razones: Que al incorporar los dos parámetros el de promedios móviles ordinarios de orden 7 y el de promedios móviles estacional de orden 1, son significativos, porque son distintos de cero estadísticamente hablando, además podemos observar de que no existe correlación significativa entre los parámetros estimados. Así en la función de autocorrelación muestral de los residuales se observa que no existe correlación entre los residuales y sus retrasos.

Considerado como un modelo adecuado podemos ver el espectro de los residuales que contiene la tabla 8.9 para definir si existe algún efecto periódico fuerte que no hayamos incorporado en el modelo.

Finalmente en la misma tabla, 8.9 podemos ver el histograma de los residuales del modelo, la desviación estándar, la media entre otros estadísticos permiten comprobar la normalidad de los residuales.

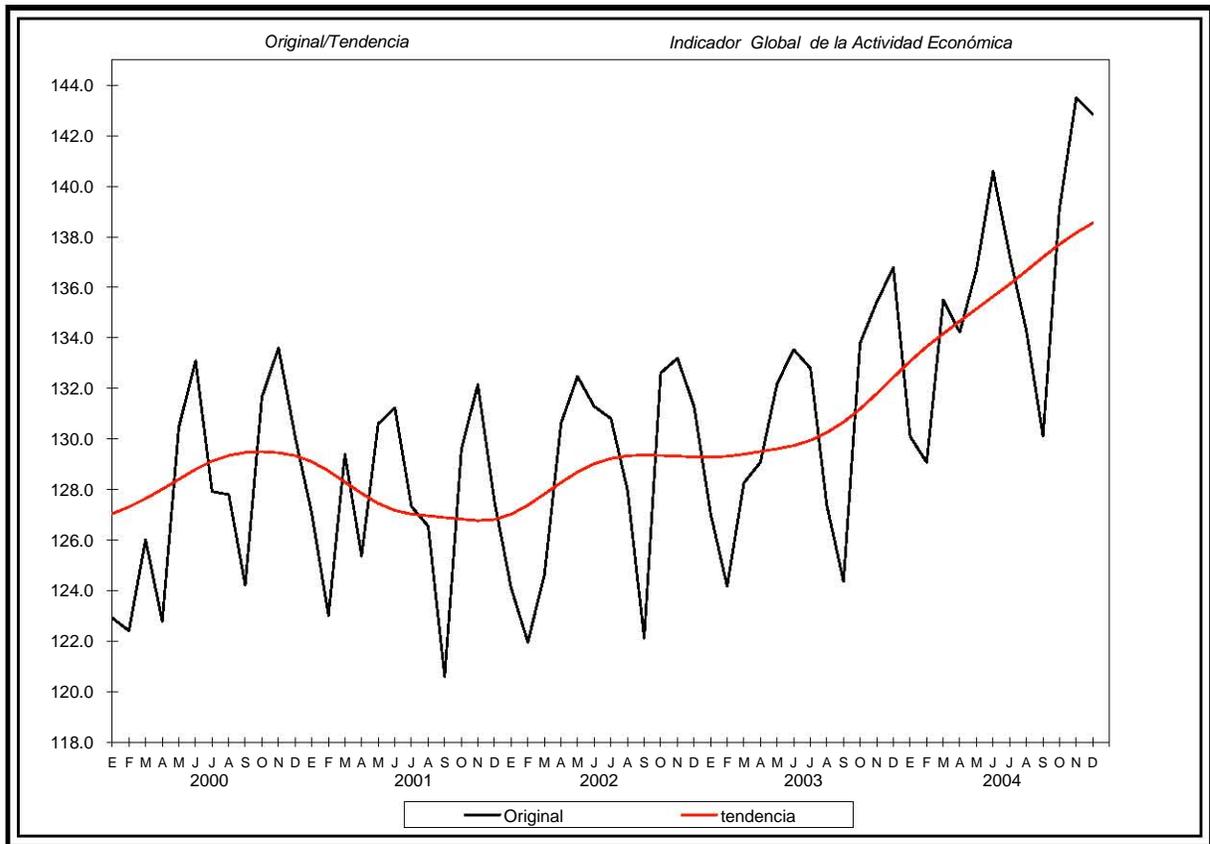
Podemos concluir que no existe algún efecto periódico conocido que hayamos dejado de incorporar al modelo, por tal razón podemos afirmar que el modelo estimado en las tablas anteriores es el más adecuado para representar las cifras desestacionalizadas del IGAE.

## 9. DESESTACIONALIZACIÓN DE LA SERIE DEL IGAE

En este apartado analizaremos de acuerdo a la metodología descrita los resultados de la serie desestacionalizada. En la ilustración siguiente se muestra la serie original así como la tendencia.

### Ilustración 9.1. México: Serie Original y Tendencia, 2000-2004.

(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Para la obtención de la tendencia se utilizó un promedio móvil centrado de 12 términos en la serie original ajustada previamente. Cabe señalar que dentro del análisis de descomposición de la serie se realizaron pruebas para detectar la presencia de estacionalidad. De estas pruebas se concluye que existe estacionalidad pero no una estacionalidad móvil en la serie, estas pruebas se presentan en la tabla 9.1.

**Tabla 9.1. México: Prueba para detectar la presencia de estacionalidad, 2000-2004.**

(Índice 1993 = 100)

Test for the presence of seasonality assuming stability.				
	Sum of Squares	Dgrs. of Freedom	Mean Square	F-Value
Between months	390.3607	11	35.48734	164.473**
Residual	10.3567	48	0.21576	
Total	400.7174	59		

\*\*Seasonality present at the 0.1 per cent level.

Nonparametric Test for the Presence of Seasonality Assuming Stability		
Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
57.1980	11	0.000%

Seasonality present at the one percent level.

Moving Seasonality Test				
	Sum of Squares	Dgrs. of Freedom	Mean Square	F-value
Between Years	1.5656	4	0.391406	1.980
Error	8.6967	44	0.197652	

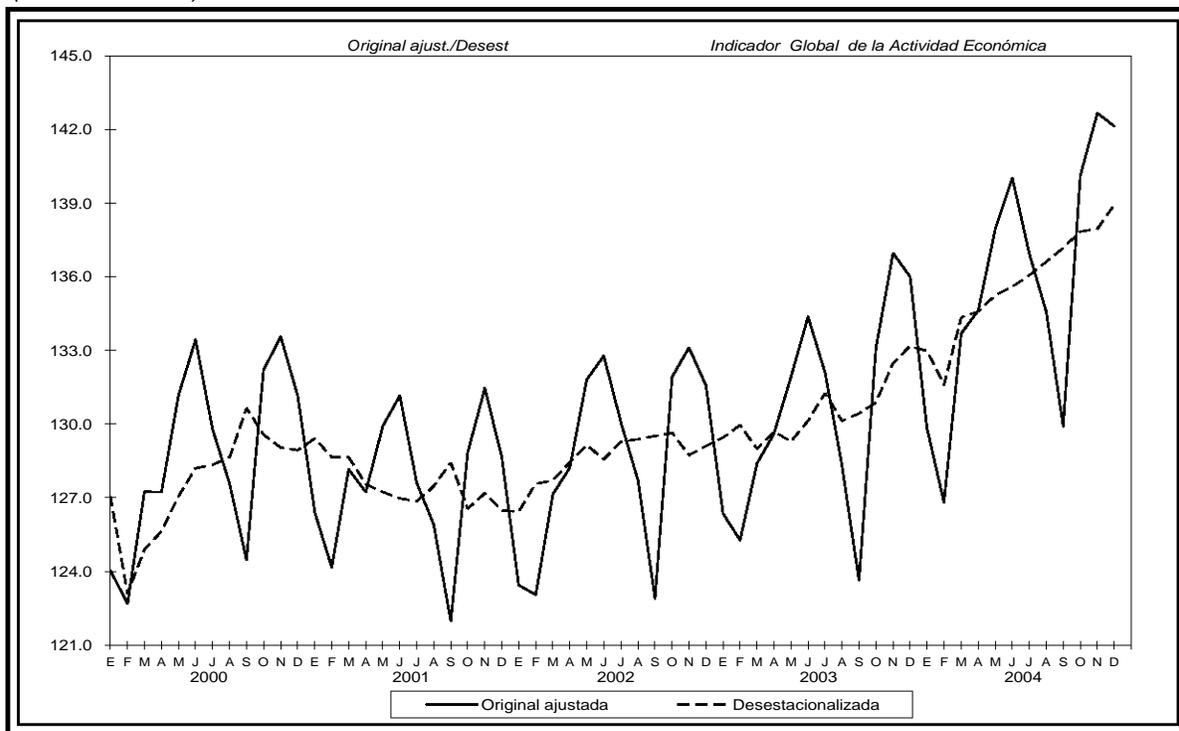
No evidence of moving seasonality at the five percent level.

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En la siguiente ilustración se presentan la serie desestacionalizada final, la serie original previamente ajustada por Trading Day, Outliers y Semana Santa. También presentamos el factor estacional final.

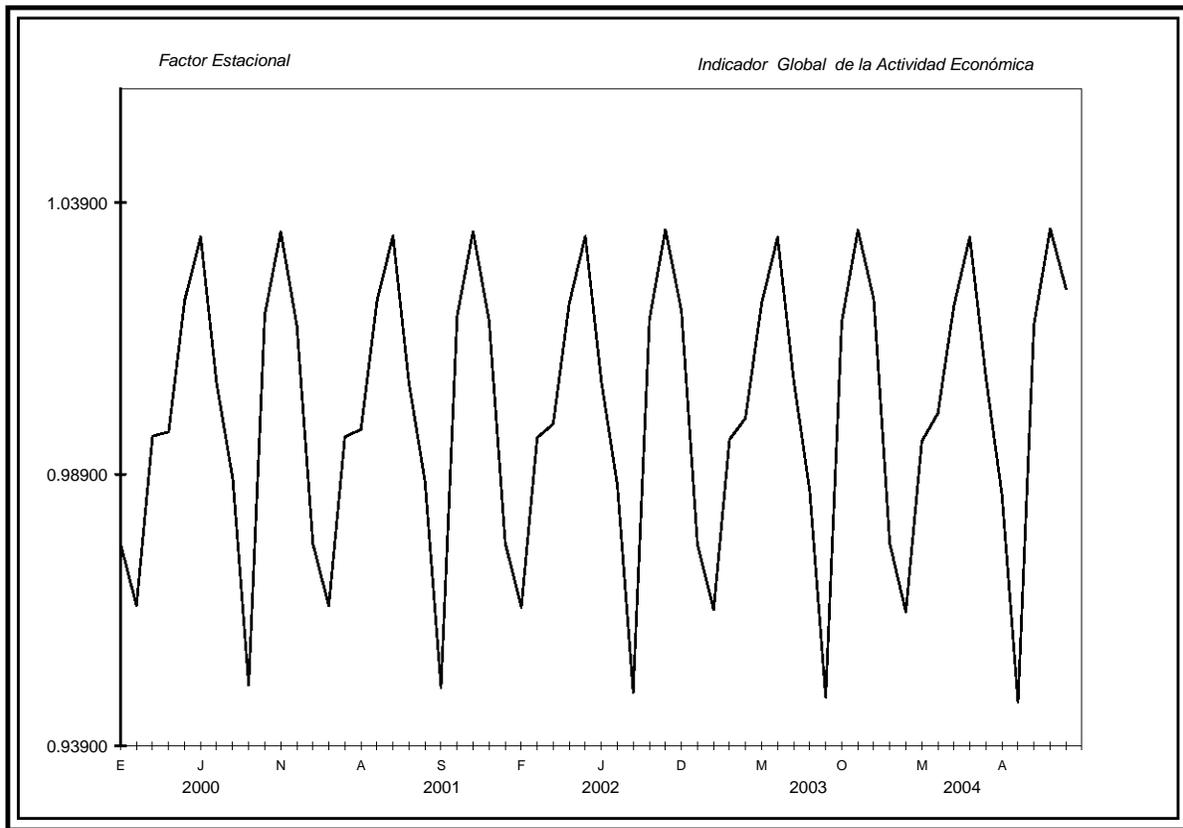
**Ilustración 9.2. México: Serie Desestacionalizada y serie original ajustada, 2000-2004.**

(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

**Ilustración 9.3. México: Factor estacional, 2000-2004.**  
(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

En la ilustración del factor estacional, se observa un comportamiento estacional muy fuerte, este comportamiento se repite en todo el periodo en estudio.

A continuación en la ilustración 9.4 se muestran las tasas de crecimiento mensual a partir de febrero de 2000 a diciembre de 2004 estas fueron calculadas de la serie desestacionalizada. Variaciones mensuales por la razón de que habiendo ajustado la serie de efectos no económicos ya podemos comparar en el margen a la serie para tener una apreciación más exacta del cambio que tenga esta.



## 10. DIAGNÓSTICO DE LA DESESTACIONALIZACIÓN

En este punto validaremos el ajuste de la serie del IGAE, estará basado en los resultados de la desestacionalización. Así el *objetivo particular será determinar si el ajuste realizado cumple con los supuestos de Box y Jenkins.*

### 10.1. ANÁLISIS DEL MODELO PARA DESESTACIONALIZAR EL IGAE

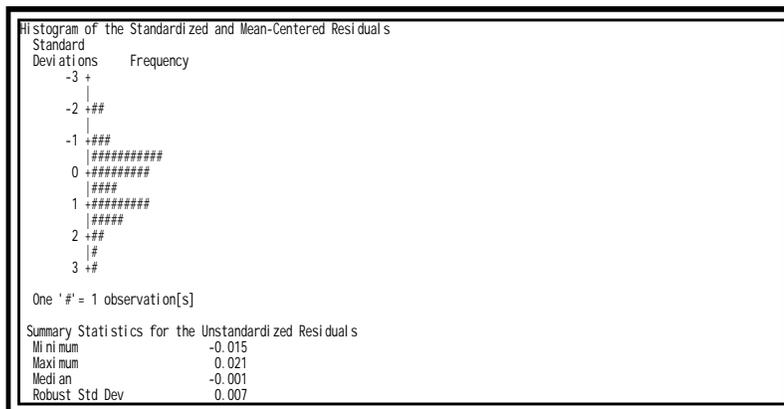
El modelo (0 1 [7]) (1 1 0) para la serie del Indicador Global de la Actividad Económica tiene las siguientes observaciones:

*En el supuesto 1 tiene media cero; efectivamente el valor absoluto -0.001 es menor que 2 no hay evidencia de que la media del proceso de ruido blanco sea distinta de cero.*

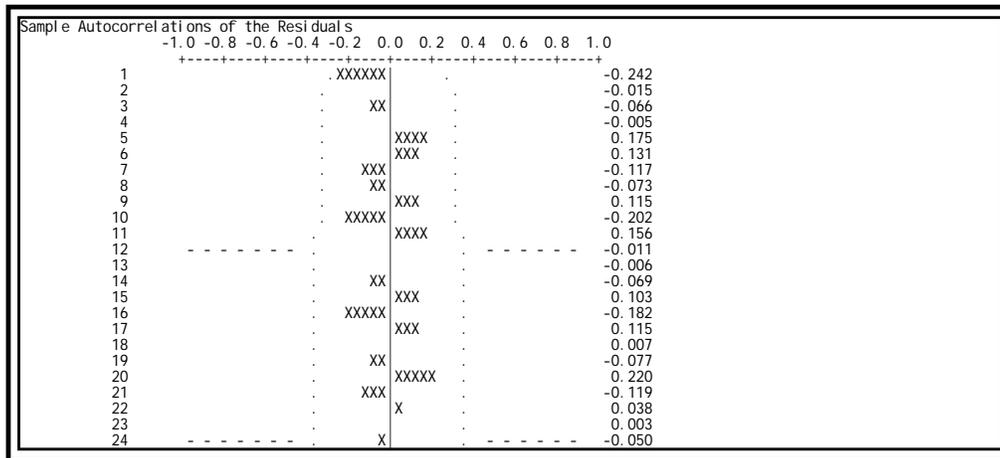
*El supuesto 2 de varianza constante, se verifica en la gráfica de residuales. Simplemente se observa que la varianza no cambia a través del tiempo.*

*En el supuesto 3 de variables aleatorias mutuamente independientes, tenemos una  $p=0.422$  y no hay datos fuera. La FAC (función de autocorrelación) es 2 veces menor que la desviación estándar. Lo podemos observar en la tabla 8.8. en las autocorrelaciones simples de los residuales.*

*El supuesto 4: Debe tener una distribución normal, en el histograma que presentamos se observa con más claridad este supuesto además de que la media tiende a cero y la desviación estándar es robusta.*



El supuesto 5 indica que: *no debe haber observaciones aberrantes*, es decir, que no estén fuera del intervalo de  $\pm 3$  desv. Estándar, este modelo cumple con este requisito al observar las autocorrelaciones simples de los residuales en la gráfica siguiente.



Además de estos supuestos *que se refieren a los residuales, se deben satisfacer los siguientes:*

El supuesto 6 considera que: *el modelo debe ser parsimonioso*, es decir que no se puede reducir el número de parámetros y no considerarse iguales a cero, ya que todos los parámetros son necesarios para explicar el fenómeno y no pueden reducirse.

El supuesto 7 considera que *el modelo es admisible*, cada parámetro tiene restricciones en los valores que puede tomar, esto puede ser constatado con el estadístico “t”, el cual podemos observar en la tabla 8.7 donde estimamos los parámetros. Con dividir el valor estimado entre el error estándar y el resultado debe ser mayor a 2.0 de lo contrario debe estimarse de nuevo el modelo. Tiene que respetarse el intervalo de confianza para que sea válido un parámetro. El modelo ARIMA (0 1 [7]) (1 1 0) cumple con los parámetros estimados.

El supuesto 8 supone que: *el modelo es estable en los parámetros y que estos en su correlación no sean mayores a 0.5*, así se debe de evitar la redundancia de parámetros es decir respetando la restricción del supuesto, por lo tanto este modelo tiene una correlación de 0.10 en sus parámetros por lo tanto no existe correlación alta.

Habiendo validado el modelo bajo los supuestos, se expone en su totalidad el modelo estimado con la finalidad de tener el resultado de todos los estadísticos encontrados; también

expresaremos la serie original y la serie desestacionalizada con sus respectivas tasas de crecimiento en forma numérica y gráfica.

Tabla 10.1. México: (1/4) modelo para desestacionalizar el IGAE, 2000-2004.

(Índice 1993 = 100)

```

C 15 Final irregular component regression
Irregular Regression Model
Trading Day + Easter[10]
Irregular Component Regression Span
From 2000.Jan to 2004.Dec

OUTLIER DETECTION
From 2000.Jan to 2004.Dec
Observations 60
Types AO only
Method add one
Critical |t| for AO outliers 3.69

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
Trading Day
Mon 0.0580 0.04100 1.41
Tue 0.1187 0.03972 2.99
Wed 0.0020 0.03901 0.05
Thu 0.0466 0.04014 1.16
Fri 0.1110 0.03909 2.84
Sat -0.0925 0.03865 -2.39
*Sun (derived) -0.2437 0.04128 -5.90
Easter[10] -0.6330 0.08371 -7.56

Automatically Identified Outliers
AO2000.Feb -0.5441 0.11358 -4.79
AO2004.Feb -0.5157 0.11413 -4.52

*For full trading-day effects, the derived parameter estimate
is obtained indirectly as minus the sum of the directly estimated
parameters that define the effect.

Chi-squared Tests for Groups of Regressors
-----
Regression Effect df Chi-Square P-Value
-----
Trading Day 6 158.46 0.00
Variance 0.11322E-01
  
```

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Tabla 10.2. México: (2/4) modelo para desestacionalizar el IGAE, 2000-2004.

(Índice 1993 = 100)

```

MODEL DEFINITION
Transformation
Log(X)
Regression Model
TC2000.02

ARIMA Model (0 1 [7]) (1 1 0)
regARIMA Model Span
From 2000.Jan to 2004.Dec

MODEL ESTIMATION/EVALUATION
Exact ARIMA likelihood estimation 200
Max total ARMA iterations 40
Max ARMA iter's w/in an IGLS iteration 1,00E-05
Convergence tolerance

NOTE: Insufficient data to compute average forecast error diagnostics.

Regression Model
-----
Variable Parameter Standard t-value
Estimate Error
-----
TC2000.02 -0.0324 0.00720 -4.50

ARIMA Model: (0 1 [7]) (1 1 0)
Nonseasonal differences: 1
Seasonal differences: 1

Parameter Value (Fixed)
-----
Seasonal AR
Lag 12 -0.4239
Nonseasonal MA
Lag 7 0.2972
Variance 0.60522E-04

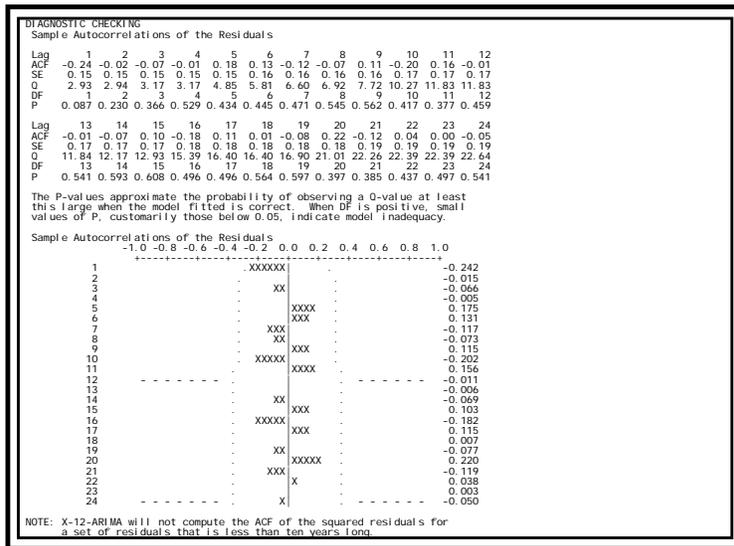
NOTE: Fixed values have been assigned to some regression and ARIMA model
coefficients. If these values are estimates calculated by X-12-ARIMA,
then the model comparison statistics (AIC, AICC, Hannan-Quinn, and BIC)
and the P-values of the Q's of the sample autocorrelations of the
residuals below are invalid and should not be used.

Likelihood Statistics
-----
Effective number of observations (nlobs) 47
Number of parameters estimated (np) 2
Log likelihood 160.0415
Transformation adjustment -229.0007
Adjusted Log likelihood (L) -68.9592
AIC 141.9164
AICC (F-corrected-AIC) 142.1911*
Hannan-Quinn 143.3108*
BIC 145.6187*

* NOTE: These statistics do not contain a penalty for parameters
estimated by x11regression to produce the prior adjustment
factors because the x11regression estimates are not maximum
likelihood estimates. Therefore they cannot be compared to
the statistics from models in which regression variables in
a regARIMA model are used to estimate the same effects.
  
```

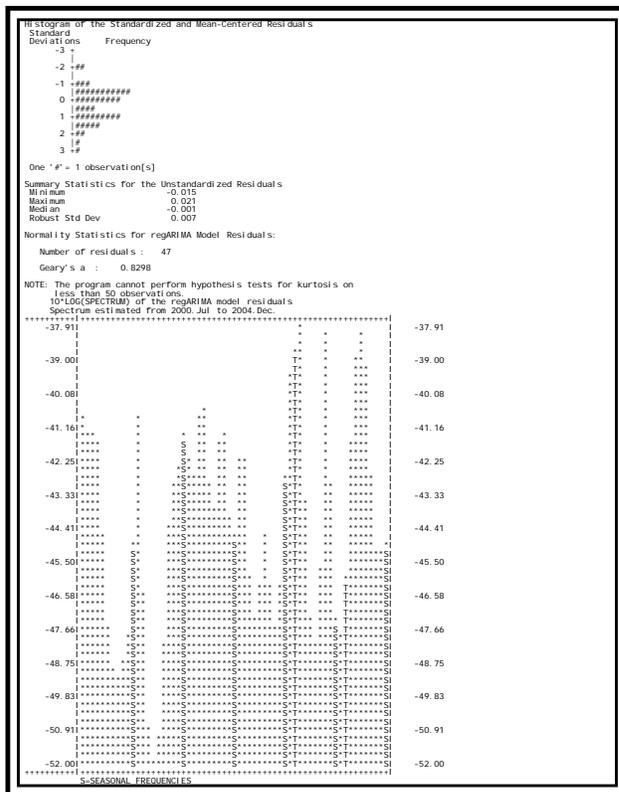
Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Tabla 10.3. México: (3/4) modelo para desestacionalizar el IGAE, 2000-2004.  
(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Tabla 10.4. México: (4/4) modelo para desestacionalizar el IGAE, 2000-2004.  
(Índice 1993 = 100)



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

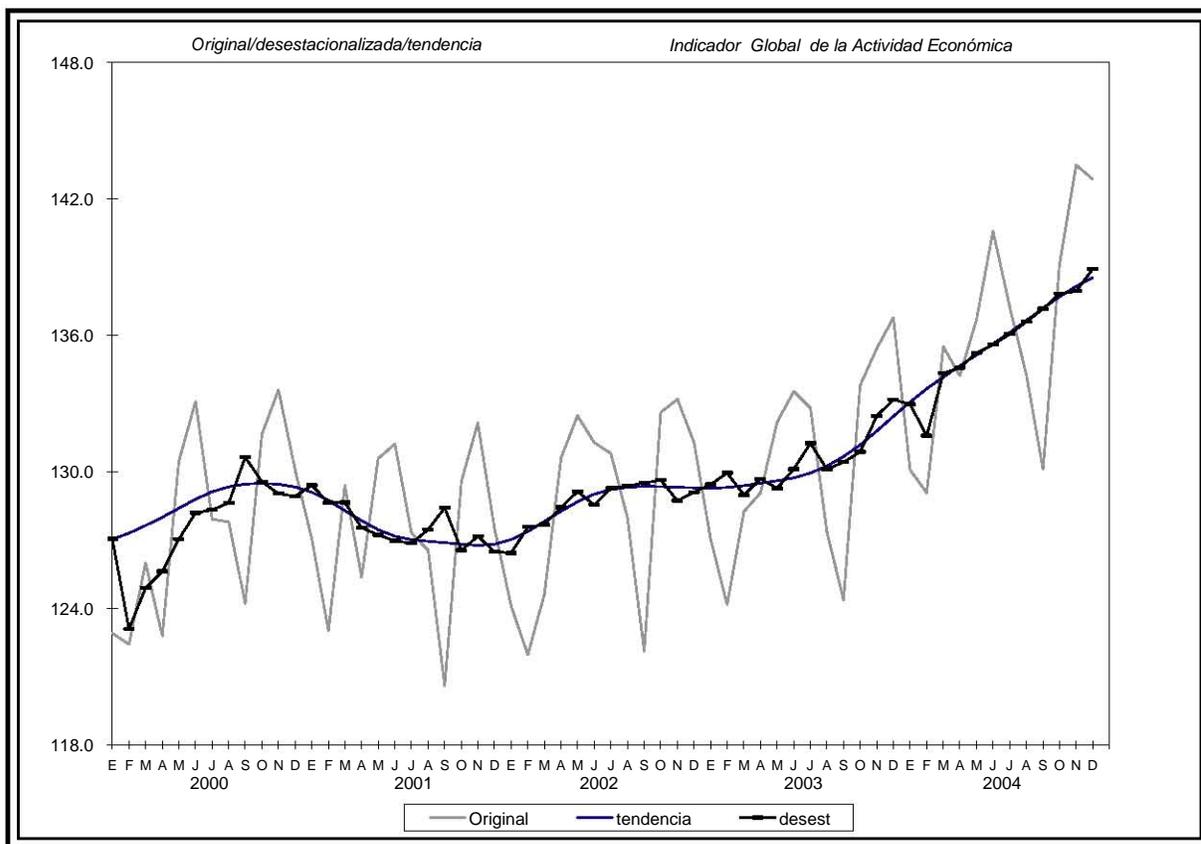
Con el modelo estimado (0 1 [7]) (1 1 0) se desestacionalizó la serie del Indicador Global de la Actividad Económica. A continuación en la tabla 10.5 se presenta la serie original, la serie desestacionalizada y la tendencia con sus tasas de crecimiento.

Tabla 10.5. México: Series original, desestacionalizada y la tendencia, 2000-2004.  
(Índice 1993 = 100)

INDICADOR GLOBAL DE LA ACTIVIDAD ECONÓMICA (IGAE) (1993=100)						
PERIODO	SERIE ORIGINAL		SERIE DESESTACIONALIZADA		TENDENCIA	
	INDICE	VAR. ANUAL	INDICE	VAR. MENSUAL	INDICE	VAR. MENSUAL
2000 E	122.9		127.1		127.0	
F	122.4		123.1	-3.12	127.3	0.22
M	126.0		124.9	1.46	127.6	0.26
A	122.8		125.6	0.59	128.0	0.29
M	130.5		127.0	1.13	128.4	0.31
J	133.1		128.2	0.91	128.8	0.31
J	127.9		128.3	0.10	129.1	0.25
A	127.8		128.6	0.24	129.3	0.17
S	124.2		130.6	1.56	129.5	0.09
O	131.7		129.6	-0.82	129.5	0.03
N	133.6		129.0	-0.40	129.5	-0.03
D	130.1		128.9	-0.09	129.3	-0.09
2001 E	127.1	3.37	129.4	0.37	129.1	-0.18
F	123.0	0.49	128.6	-0.59	128.7	-0.28
M	129.4	2.68	128.6	0.00	128.3	-0.35
A	125.4	2.10	127.5	-0.85	127.8	-0.35
M	130.6	0.10	127.2	-0.25	127.5	-0.31
J	131.2	-1.40	127.0	-0.21	127.2	-0.22
J	127.3	-0.47	126.9	-0.09	127.0	-0.12
A	126.6	-0.97	127.5	0.48	127.0	-0.06
S	120.6	-2.91	128.4	0.75	126.9	-0.05
O	129.6	-1.58	126.6	-1.46	126.8	-0.05
N	132.1	-1.08	127.2	0.49	126.8	-0.05
D	127.5	-1.97	126.5	-0.54	126.8	0.04
2002 E	124.1	-2.32	126.4	-0.04	127.0	0.16
F	122.0	-0.86	127.6	0.90	127.4	0.28
M	124.6	-3.69	127.7	0.09	127.8	0.35
A	130.6	4.19	128.4	0.59	128.3	0.35
M	132.5	1.44	129.1	0.55	128.7	0.32
J	131.3	0.05	128.6	-0.45	129.0	0.26
J	130.8	2.74	129.3	0.56	129.2	0.16
A	128.0	1.12	129.4	0.08	129.3	0.08
S	122.1	1.26	129.5	0.10	129.4	0.03
O	132.6	2.34	129.6	0.10	129.3	-0.01
N	133.2	0.79	128.7	-0.70	129.3	-0.02
D	131.3	2.96	129.1	0.29	129.3	-0.02
2003 E	127.0	2.34	129.4	0.26	129.3	-0.01
F	124.2	1.81	129.9	0.39	129.3	0.03
M	128.3	2.92	129.0	-0.74	129.4	0.06
A	129.1	-1.19	129.7	0.55	129.5	0.08
M	132.2	-0.21	129.3	-0.32	129.6	0.08
J	133.5	1.71	130.1	0.65	129.7	0.10
J	132.8	1.52	131.3	0.87	129.9	0.16
A	127.4	-0.46	130.1	-0.86	130.3	0.24
S	124.4	1.83	130.4	0.23	130.7	0.32
O	133.8	0.89	130.9	0.35	131.2	0.40
N	135.4	1.68	132.5	1.21	131.8	0.46
D	136.8	4.19	133.2	0.54	132.4	0.49
2004 E	130.1	2.43	133.0	-0.15	133.1	0.48
F	129.1	3.95	131.6	-1.04	133.7	0.43
M	135.5	5.65	134.3	2.10	134.2	0.39
A	134.2	3.99	134.6	0.17	134.7	0.36
M	136.7	3.40	135.2	0.49	135.1	0.36
J	140.6	5.28	135.6	0.26	135.6	0.36
J	137.2	3.35	136.1	0.34	136.1	0.36
A	134.3	5.42	136.6	0.40	136.7	0.39
S	130.1	4.62	137.2	0.41	137.2	0.39
O	139.1	3.96	137.8	0.49	137.7	0.37
N	143.5	5.96	137.9	0.07	138.2	0.33
D	142.9	4.44	138.9	0.71	138.5	0.28

Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

**Ilustración 10.1. México: Serie original, desestacionalizada y la tendencia, 2000-20005.**  
*(Índice 1993 = 100)*



Fuente: Elaboración propia con datos del Instituto Nacional de Estadística, Geografía e Informática (INEGI).

Al observar la anterior tabla con su gráfica, podemos afirmar que los efectos de estacionalidad presentes en la serie original fueron eliminados por el proceso de desestacionalización; también podemos argumentar que a partir de que fue desestacionalizada la serie del IGAE ya podemos calcular tasas de crecimiento mensuales para un análisis más exacto de ésta.

## 11. CONCLUSIONES Y RECOMENDACIONES

La utilización de datos estadísticos y particularmente de datos económicos, tienen su importancia cada vez en mayor medida para los diferentes sectores sociales, es por ello que es trascendental que estos datos sean en importancia relevante y cuidadosamente elaborados por las instituciones dedicadas a difundirlos.

Como se mencionó en el presente trabajo, la serie del Indicador Global de la Actividad Económica era afectada por diversos factores ajenos a ésta como la presencia de Semana Santa, los días laborables o valores extremos entre otros factores, que de alguna forma distorsionan el análisis. Como ejemplo, podemos mencionar que una comparación propiamente de este indicador en su serie original no podría hacerse mes con mes por estar influenciada por factores estacionales.

Podemos dar una infinidad de ejemplos donde las series no han sido ajustadas estacionalmente lo cual nos conduciría a conclusiones erróneas.

Así, con anterioridad se mencionó que existen diferentes efectos contenidos en una serie económica, y a la vez mecanismos para corregirlos o aislarlos de las series. Para corregir la estacionalidad, objetivo principal de esta tesina y que está presente en la mayoría de las series económicas; y que pueda este factor ser disminuido en su máxima medida, con el objetivo de hacer un análisis más puro.

Para lograr lo anterior fue necesario proponer la utilización del método X12-ARIMA, desarrollado por el U.S. Census Bureau, este método es una derivación de los métodos X11 y X11-ARIMA de una amplia difusión. Como se mencionó anteriormente no es el único método que se utiliza para desestacionalizar series económicas. Con lo anterior no se pretende argumentar que el método descrito es el mejor para lograr el ajuste estacional, sin embargo si se pretende mostrar las bondades que proporciona el método para la manipulación de los datos de la economía mexicana.

Es importante mencionar que puede hallarse otra metodología para el análisis; pero se utilizó la que mejor se adapta a los datos de la economía mexicana, de esta forma es importante mencionar que un conocimiento sólido de algún método de ajuste de series económicas nos ayuda a un análisis más adecuado.

Por otra parte, aunque ya se mencionaron las bondades del método utilizado, es necesario establecer las limitaciones de este método, principalmente cuando se hacen conclusiones de un análisis de coyuntura. Esto es que para lograr un ajuste en los datos finales de la serie se hace uso de la metodología de Box y Jenkins para establecer un modelo ARIMA para concretar las técnicas de desestacionalización en una serie. Cabe destacar que el estudio sobre series de tiempo desarrollado por estos autores, es un análisis univariado que no previene los cambios en las variables económicas. Es importante resaltar que cuando se van agregando más datos, los ajustes se van haciendo más robustos para los datos intermedios.

Otra característica importante de este trabajo es que los datos utilizados fueron para una serie económica real de cinco años; sin embargo existe una serie más amplia de este indicador mensual desestacionalizada base 1993, en el Banco de Información Económica del INEGI.

Por lo antes dicho podemos afirmar que la metodología descrita y aplicada a la serie del Indicador Global de la Actividad Económica para ajustarlo resulta eficaz.

## 12. ANEXO: SUPUESTOS (BOX-JENKINS)

*Supuesto 1.  $\{a_t\}$  tiene media cero*

Verificación: Calcular el promedio aritmético y la desviación estándar muestral de los residuales estimados.

(9)	$m(\hat{a}) = \frac{\sum_{t=t'}^N \hat{a}_t}{(N-d-p)}$
(10)	$\hat{\sigma}_a = \sqrt{\frac{\sum_{t=t'}^N [\hat{a}_t - m(\hat{a})]^2}{(N-d-p-q)}}$

donde  $t' = d + p + 1$

y calcular el cociente

$\frac{\sqrt{N-d-p} m(\hat{a})}{\hat{\sigma}_a}$
--

Si el valor absoluto de este cociente es menor que dos, entonces no existe evidencia de que la media de proceso de ruido blanco sea distinta de cero. Por el contrario, si es mayor a dos, la media de los residuales es significativamente distinta de cero, por lo que el supuesto ha sido violado.

Corrección: Si este supuesto no se satisface, quiere decir que existe una parte determinista o semideterminista en  $\{a_t\}$  que no fue considerado en el modelo, por lo tanto es necesario incluir una tendencia determinista (constante) que deberá ser estimada conjuntamente con el resto de los parámetros.

*Supuesto 2.  $\{a_t\}$  tiene varianza constante*

Verificación: la prueba de la validez de este supuesto es visual. Se deben graficar los residuales estimados contra el tiempo y observar si la varianza parece o no ser constante.

Corrección: Si la varianza parece seguir algún patrón creciente o decreciente, se debe verificar la transformación potencia que estabilice la varianza.

*Supuesto 3.  $\{a_t\}$  son mutuamente independientes*

Verificación: ya que la independencia implica no-auto correlación, entonces la auto correlación del retraso  $k$  de  $\{a_t\}$  debe ser cero para toda  $k \neq 0$ . Esto último se verifica calculando la función de auto correlación muestral de los residuales  $\{r_k(\hat{a})\}$ , que bajo el supuesto de media cero, ésta se define como:

$$(11) \quad r_k(\hat{a}) = \frac{\sum_{t=t'}^{N-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=t'}^N \hat{a}_t^2}, \quad k = 1, 2, \dots$$

Para determinar la significancia estadística individual de las autocorrelaciones de los residuales, se estima la desviación estándar de (11), la cual es aproximadamente.

$$\frac{1}{\sqrt{N-d-p}}$$

Si el valor absoluto de la  $k$ -ésima auto correlación es mayor a dos veces la desviación estándar de la función de auto correlación muestral de los residuales, se rechazará la hipótesis nula de que  $r_k(\hat{a}) = 0$ . Esta prueba de significancia individual no es válida para los primeros retrasos ( $k \leq 3$ ), por lo que conviene hacer una prueba conjunta de la significancia estadística de las primeras  $K$  auto correlaciones.

La hipótesis nula es que las primeras  $K$  auto correlaciones muestrales de los residuales son igual a cero. Se calcula el estadístico  $Q$  de Ljung y Box donde:

$$(12) \quad Q = \frac{(N-d-p)(N-d-p+2) \sum_{k=1}^K r_k^2(\hat{a})}{(N-d-p-k)},$$

El cual sigue una distribución  $\chi^2$  con  $K-p-q$  grados libertad. Por lo tanto, se rechaza la hipótesis nula si  $Q$  es mayor a una  $\chi^2(K-p-q, \alpha)$ , donde  $\alpha$  es el nivel de significancia estadística. En la práctica el valor de  $K$  se fija en 24, por lo que las primeras 24 auto correlaciones deben ser iguales a cero.

*Corrección:* Si la verificación indica que las auto correlaciones no son las correspondientes a un proceso de ruido blanco, se deberá identificar algún proceso ARMA

para los residuales en  $\{r_k(\hat{a})\}$ , lo cual sugerirá modificaciones al modelo originalmente identificado para la serie.

*Supuesto 4.  $a_t$  tiene una distribución normal, para toda  $t$*

*Verificación:* la verificación de este supuesto se realiza con dos pruebas distintas. La primera es la prueba de Jarque y Bera, la segunda se observa si el 95% de los residuales estandarizados se encuentra dentro del intervalo  $(-2\sigma_a, 2\sigma_a)$ .

*Corrección:* la corrección de este supuesto es difícil, pero cabe señalar que este supuesto debe cumplirse de manera exacta para  $\{a_t\}$  y no para  $\{\hat{a}_t\}$ . Por lo que violaciones pequeñas pueden esperarse: si las violaciones son grandes, es posible aplicar alguna transformación del tipo de Box-Cox<sup>19</sup>. En la práctica se evita usar transformaciones normalizantes, pues al aplicarlas se pierde interpretación de las variables.

*Supuesto 5. No existen observaciones aberrantes*

*Verificación:* al igual que en el supuesto anterior se ve si existe algún residual que se encuentre fuera del intervalo  $(-3\sigma_a, 3\sigma_a)$ .

*Supuesto 6. El modelo es parsimonioso*

*Verificación:* por parsimonia se entiende que no se puede reducir el número de parámetros involucrados, ya que todos son necesarios para explicar el comportamiento de la serie y no pueden ser considerados iguales a cero. Lo que se hace es calcular intervalos de confianza de aproximadamente 95% de confianza para cada uno de los parámetros y observar si el valor de cero pertenece al intervalo.

$$\left( \hat{\theta} - 2\sqrt{\text{Var}(\hat{\theta})}, \hat{\theta} + 2\sqrt{\text{Var}(\hat{\theta})} \right).$$

*Corrección:* Sí el cero es un valor razonable para algún parámetro, entonces se debe omitir y volver a estimar el modelo sin él. Sin embargo, pueden existir situaciones en las que el conocimiento del fenómeno indique que el parámetro debe aparecer, aun cuando la verificación no rechace la hipótesis nula de que su valor sea cero.

---

<sup>19</sup> Para mayor información de la transformación del tipo de Box y Cox, consultar la siguiente página: <http://www.ucm.es/info/ecocuan/mjm/ectr2mj/series2.pdf>

Debido a que la función de verosimilitud de un modelo ARIMA es no lineal, la maximización de ésta no puede obtenerse analíticamente, por lo que son necesarios métodos numéricos para resolver el problema de maximización, por lo cual usamos los estimadores de máximo verosímiles del modelo ARIMA del algoritmo de Gauss-Newton.

*Supuesto 7. El modelo es admisible*

Cada parámetro tiene restricciones en los valores que puede tomar. Una violación a éstas hace que no se pueda aceptar el modelo.

*Supuesto 8. El modelo es estable en los parámetros*

La principal causa de inestabilidad se debe a la redundancia de parámetros, es decir, que un parámetro puede ser explicado por otros. Esto se puede observar a través de las correlaciones entre parámetros, donde una alta correlación podría sugerir un parámetro inestable.

## BIBLIOGRAFÍA

1. Baena Guillermina; Sergio Montero; Tesis en 30 días, 13ª. Reimpresión, Ed. Mexicanos Unidos, 1996.
2. Box, George E. P., G. W. Jenkins y G. C. Reinsel (1994) *Time series analysis: forecasting and control*, Nueva Jersey, Prentice & Hall (3ª edición).
3. CEPAL - SERIES Manuales, No. 9. Cuarta parte capítulos 8 y 9.
4. Dagum, Estela Bee (1978) *A comparison and assessment of seasonal adjustment methods for employment and unemployment statistics*, Washington, DC. Background Paper No. 7, National Commission on Employment and Unemployment Statistics.
5. Dominique Ladiray, Benoît Quenneville “Desestacionalizar con el Método X-11”; Prefacio de Allan Young; Traducción al castellano: Eduardo Crivisqui 2000 – 2001, Documento en PDF.
6. Findley, David F. *et al.* (1998) “New capabilities and methods of the X-12-ARIMA seasonal adjustment program,” *Journal of Business and Economic Statistics*, 16(2): 127-177.
7. Guerrero, Víctor (1983a) “Desestacionalización de series de tiempo económicas: parte I. Una introducción a la metodología”, documento de investigación No. 54 Banco de México.
8. Guerrero, Víctor (1983b) “Desestacionalización de series de tiempo económicas”, parte II Ajustes previos a la desestacionalización, documento de investigación No. 55 Banco de México.
9. Guerrero, Víctor; (1991) *Análisis estadístico de series de tiempo económicas*, México, D.F. UAM Iztapalapa, (Colección cbi).
10. Guerrero, Víctor; Rojas Javier; (1987) “Desestacionalización de series de tiempo económicas”, Aplicación a los Indicadores de la Actividad Industrial, documento de investigación No. 63 Banco de México.
11. Heath Jonathan; Ordóñez Ivonne; Desaceleración, Recesión o Crisis 2001, en la CIV Junta Macroeconómica de CIEMEX-WEFA.
12. Hernández Sampieri Roberto; Fernández Collado Carlos; Pilar Baptista Lucio; Metodología de la Investigación; 1ª ed. México 1991. Ed. McGraw Hill.
13. <http://www.mecon.gov.ar/secpro/dir-cn/desestacionalización/doc0.htm>
14. <http://www.uam.es/departamentos/económicas/econapli/pdf/trata-ser.pps>
15. <http://www.ucm.es/info/ecocuan/mjm/ectr2mj/series2.pdf>
16. INEGI; (2000) “Sistema de Cuentas Nacionales de México”, Indicador Global de la Actividad Económica (Mensual).
17. Núñez del Prado Benavente Arturo; estadística Básica para Planificación; 16ª. Ed. Siglo Veintiuno.

18. Pardinas Felipe; Metodología y Técnicas de Investigación en Ciencias Sociales; 37ª. Ed. Siglo Veintiuno.
19. Pindyck Robert S; Rubinfeld Daniel L; Econometría, Modelos y Pronósticos; 4a. ed. México, 2001, Ed. McGraw Hill.
20. U.S. Census Bureau; X-12-ARIMA, Seasonal Adjustment Program.
21. Vuskovic Bravo, Pedro. Los Instrumentos Estadísticos del Análisis Económico. Cide 1988, México, 1986.
22. Zorrilla Santiago; Torres Miguel; Guía para Elaborar la Tesis Ed. McGRAW-HILL; 1990.

## ÍNDICE DE TABLAS

<i>Tabla 3.1. México: Cálculos para seleccionar la potencia (<math>\lambda</math>) de la Transformación estabilizadora de la varianza.</i>	15
<i>Tabla 4.1. Decisión entre diferentes filtros de Henderson.</i>	22
<i>Tabla 7.1. México: Cálculo para seleccionar la potencia de la transformación estabilizadora de la varianza del IGAE, 2000-2004.</i>	33
<i>Tabla 7.2. México: Espectro de la serie original transformada, 2000-2004.</i>	36
<i>Tabla 7.3. México: Td y Easter, 2000-2004.</i>	37
<i>Tabla 7.4. México: Serie original corregida por días laborables y semana santa, 2000-2004.</i>	39
<i>Tabla 7.5. México: Identificación de outliers, 2000-2004.</i>	40
<i>Tabla 7.6. México: Diferencia estacional (0 1) de la serie transformada y ajustada por semana santa, trading day y outliers, 2000-2004.</i>	42
<i>Tabla 7.7. México: Diferencia no estacional (1 0) de la serie transformada y ajustada por semana santa, trading day y outliers, 2000-2004.</i>	43
<i>Tabla 7.8. México: Diferencia no estacional y estacional (1 1) de la serie transformada y ajustada por semana santa, trading day y outliers, 2000-2004.</i>	44
<i>Tabla 8.1. México: Estimación de los parámetros para el modelo (0 1 0) (0 1 1), 2000-2004.</i>	45
<i>Tabla 8.2. México: Correlación de los parámetros para el modelo (0 1 0) (0 1 1), 2000-2004.</i>	46
<i>Tabla 8.3. México: Autocorrelación de los residuales para el modelo (0 1 0) (0 1 1), 2000-2004.</i>	47
<i>Tabla 8.4. México: Espectro de los residuales para el modelo (0 1 0) (0 1 1), 2000-2004.</i>	48
<i>Tabla 8.5. México: estimación del modelo (0 1 0) (1 1 0), 2000-2004.</i>	49
<i>Tabla 8.6. México: diagnóstico del modelo (0 1 0) (1 1 0), 2000-2004.</i>	50
<i>Tabla 8.7. México: estimación de los parámetros para el modelo (0 1 [7]) (1 1 0), 2000-2004.</i>	51
<i>Tabla 8.8. México: parámetros estimados para el modelo (0 1 [7]) (1 1 0), 2000-2004.</i>	52
<i>Tabla 8.9. México: histograma y espectro del modelo (0 1 [7]) (1 1 0), 2000-2004.</i>	53
<i>Tabla 9.1. México: Prueba para detectar la presencia de estacionalidad, 2000-2004.</i>	56
<i>Tabla 10.1. México: (1/4) modelo para desestacionalizar el IGAE, 2000-2004.</i>	61
<i>Tabla 10.2. México: (2/4) modelo para desestacionalizar el IGAE, 2000-2004.</i>	61
<i>Tabla 10.3. México: (3/4) modelo para desestacionalizar el IGAE, 2000-2004.</i>	62
<i>Tabla 10.4. México: (4/4) modelo para desestacionalizar el IGAE, 2000-2004.</i>	62
<i>Tabla 10.5. México: Series original, desestacionalizada y la tendencia, 2000-2004.</i>	63

## ÍNDICE DE ILUSTRACIONES

<i>Ilustración 6.1. México: Composición del PIB por Grandes Divisiones.</i>	28
<i>Ilustración 6.2. México: Estructura Porcentual del IGAE.</i>	30
<i>Ilustración 6.3. México: Sectores y Divisiones que Componen el IGAE.</i>	31
<i>Ilustración 6.4. México: Diagrama de ajuste del IGAE.</i>	32
<i>Ilustración 7.1. México: IGAE serie original, 2000-2004.</i>	35
<i>Ilustración 7.2. México: factores combinados semana santa/días laborables 2000-2004.</i>	38
<i>Ilustración 7.3. México: Serie original y la serie ajustada por semana santa y efectos de calendario 2000-2004.</i>	39
<i>Ilustración 7.4. México: Serie original y la serie corregida por semana santa, Trading day y outliers, 2000-2005.</i>	41
<i>Ilustración 9.1. México: Serie Original y Tendencia, 2000-2004.</i>	55
<i>Ilustración 9.2. México: Serie Desestacionalizada y serie original ajustada, 2000-2004.</i>	56
<i>Ilustración 9.3. México: Factor estacional, 2000-2004.</i>	57
<i>Ilustración 9.4. México: Tasas de crecimiento mensual de la serie desestacionalizada, 2000-2004.</i>	58
<i>Ilustración 10.1. México: Serie original, desestacionalizada y la tendencia, 2000-2005.</i>	64